Journal of the
AMERICAN STATISTICAL ASSOCIATION

JUNE 1958

Measuring Recessions ..... *Geoffrey H. Moore* 289
The Accuracy and Structure of Industry Expectations in Relation to Those of Individual Firms ..... *Robert Ferber* 317
An Index of Manufacturing Production in New England ..... *Harry Benjamin Ernst* 336
Effect of Varying Degrees of Transitory Income on Income Elasticity of Expenditures ..... *Marilyn Dunsing and Margaret G. Reid* 348
The First 1,945 British Steamships ..... *J. R. T. Hughes and Stanley Reiter* 360
The Use of Random Work Sampling for Cost Analysis and Control ..... *A. C. Rosander, H. E. Guterman, and A. J. McKeon* 382
The Trentile Deviation Method of Weather Forecast Evaluation ..... *Morris James Slonim* 398
Weight-Height Standards Based on World War II Experience ..... *Bernard D. Karpinos* 408
Actuarial Estimation of Survivorship in Chronic Disease ..... *Lila Elveback* 420
Graphic Computation of Tau as a Coefficient of Disarray ..... *Harold D. Griffin* 441
A New Bivariate Sign Test ..... *Isadore Blumen* 448
Nonparametric Estimation from Incomplete Observations ..... *E. L. Kaplan and Paul Meier* 457
Inadmissible Samples and Confidence Limits ..... *Howard L. Jones* 482
The Precision of Unbiased Ratio-Type Estimators ..... *Leo A. Goodman and H. O. Hartley* 491
On Noncoverage of Sample Dwellings ..... *Leslie Kish and Irene Hess* 509
Complete Counterbalancing of Immediate Sequential Effects in a Latin Square Design ..... *James V. Bradley* 525
Factorial Experimentation in Scheffé's Analysis of Variance for Paired Comparisons ..... *Otto Dykstra, Jr.* 529
Design and Operation of a Double-Limit Variables Sampling Plan ..... *Acheson J. Duncan* 543
Empiric Investigation of a Test of Homogeneity for Populations Composed of Normal Distributions ..... *G. A. Baker* 551
NOTES ABOUT AUTHORS 558
SUMMARIES OF PAPERS DELIVERED AT 117TH ANNUAL MEETING .. 565
BOOK REVIEWS 595
PUBLICATIONS RECEIVED 628
REPORT OF BOARD OF DIRECTORS, 1957 629

VOLUME 53 NUMBER 282

American Statistical Association

Organized November 27, 1839, Incorporated 1841

The American Statistical Association is a scientific and educational organization. Its membership is not confined to professional statisticians but includes economists, business executives, research directors, government officials, university professors, and others who are seriously interested in the application of statistical methods to practical problems, in the development of more useful methods, and in the improvement of basic statistical data. Engineers, mathematicians, biologists, actuaries, sociologists, psychologists, and representatives of many other professions are included in the membership of the Association.

ANNUAL DUES*

Residents of North America	$ 8	Institutional membership, minimum	160
Others	5	Member's subscription to *Biometrics*	4
Student membership	5	Member's subscription to *The American Sociological Review*	5
Introductory membership (first dues payment of members under 30 years old)	5		

* Includes subscription to the *Journal of the American Statistical Association* and to *The American Statistician*.

Subscription rate, $8 per year. Prices for back issues available on request.

For further information about the Association and membership application forms, write the Secretary, 1757 K St., N.W., Washington 6, D. C.

The present Institutional Members are:

American Telephone & Telegraph Company
Armour & Company
Armstrong Cork Company
The Atlantic Refining Company
Bankers Trust Company
Bell Telephone Laboratories
California Texas Oil Company, Ltd.
Chrysler Corporation
Deere & Company
Dun and Bradstreet, Inc.
Eastman Kodak Company
The Equitable Life Assurance Society of the U. S.
Ford Motor Company
General Dynamics Corporation
General Foods Corporation
General Motors Corporation
Humble Oil and Refining Co.
International Brotherhood of Teamsters, Chauffeurs, Warehousemen & Helpers of America
International Business Machines Corp.
International Harvester Company
International Ladies Garment Worker's Union
Metropolitan Life Insurance Company
National Analysts, Inc.
National Association of Manufacturers
National Industrial Conference Board, Inc.
New York Telephone Company
Philadelphia Electric Company
Schering Corporation
Sears Roebuck & Company
Socony Mobil Oil Company, Inc.
Standard Oil Company (Indiana)
Standard Oil Company of New Jersey
Statistical Tabulating Corporation
Sun Oil Company
Touche, Niven, Bailey & Smart
Union Carbide Corporation
United States Steel Corporation
The Wellington Fund, Inc.
Western Electric Company, Inc.
Western Union Telegraph Company
Young & Rubicam, Inc.

Published Quarterly by the AMERICAN STATISTICAL ASSOCIATION

Publication Office: 450 Ahnaip Street, Menasha, Wisconsin. Editorial Office: 116 Haskell Hall, University of Chicago, Chicago 37, Illinois. Business Office: 1757 K St., N.W., Washington 6, D. C. Acceptance for mailing at special rates of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, section 538, P. L. & R., authorized March 28, 1928. Entered as second class matter at the post office at Menasha, Wisconsin.

Anyone wishing to change his mailing address should allow eight weeks notice. A copy of the address taken from an issue of the periodical should accompany the change of address request.

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

The Editors welcome the submission of manuscripts for possible publication. They should be typewritten entirely double-spaced, including footnotes, and two copies should be sent to the Editor, W. Allen Wallis, 114 Haskell Hall, University of Chicago, Chicago 37. Books for review should be sent to the same address. Unsolicited book reviews are not accepted, but suggestions of titles for review are welcome.

EDITOR

W. ALLEN WALLIS, *University of Chicago*

ASSISTANT TO THE EDITOR: CAROLYN HIGINBOTHAM

ASSOCIATE EDITORS

CHURCHILL EISENHART
National Bureau of Standards

MILLARD HASTAY
City College of New York

CLIFFORD HILDRETH
Michigan State University

WILLIAM G. MADOW
Stanford Research Institute

I. RICHARD SAVAGE
University of Minnesota

ADVISORY PANEL OF FORMER EDITORS

WILLIAM G. COCHRAN (1945–50)
Harvard University

FRANK A. ROSS (1926–34, 41–45)
Thetford, Vermont

WILLIAM F. OGBURN (1920–25)
University of Florida

FREDERICK F. STEPHAN (1935–40)
Princeton University

LIAISON EDITORS FROM SECTIONS

BIOMETRICS: Jerome Cornfield, *National Institutes of Health, Bethesda, Maryland*

BUSINESS AND ECONOMIC STATISTICS: Frank Garfield, *Federal Reserve Board, Washington 25, D. C.*

PHYSICAL AND ENGINEERING SCIENCES: Irving Burr, *Purdue University, Lafayette, Indiana*

SOCIAL STATISTICS, Herbert Solomon, *Teachers College, Columbia University, New York*

TRAINING OF STATISTICIANS: John S. Freund, *Arizona State College, Tempe, Arizona*

***Corrigenda:** Readers and authors are urged to submit to the Editor notices of errors found in this or any previous volume. These will be published once a year, in the December issue.*

WILEY BOOKS

SAMPLING OPINIONS

An Analysis of Survey Procedure

By FREDERICK F. STEPHAN, *Princeton University; and* PHILIP J. McCARTHY, *Cornell University*. Designed to provide the reader with a better understanding of sampling in order that he may gauge its dependability and use its results intelligently. It treats survey design in broad nonrestrictive terms, discussing both sampling and the research method. The authors examine the complex procedures that are required for large scale surveys, including problems that arise in ordinary studies to which sampling theory has not been applied before. One of the WILEY PUBLICATIONS IN STATISTICS, Walter A. Shewhart and S. S. Wilks, Editors. *1958. 451 pages. $12.00.*

INTRODUCTION to DIFFERENCE EQUATIONS

With Illustrative Examples from Economics, Psychology, and Sociology

By SAMUEL GOLDBERG, *Oberlin College*. Reflects the growing interest throughout the country in the application of mathematics to the social sciences. The book offers a comprehensive introduction to the principles and applications of finite differences and difference equations. The author includes examples from the literature of economics, psychology, and sociology as a means of demonstrating the various ways difference equations can be employed outside the field of mathematics. *1958. 260 pages. $6.75.*

EXPERIMENTAL DESIGNS in INDUSTRY

Edited by VICTOR CHEW, *North Carolina State College*. Examines the applications of experimental design and statistical techniques to industrial research. The book includes a consideration of fractional factorials, response surface designs, and regression analyses. One of the WILEY PUBLICATIONS IN STATISTICS, Walter A. Shewhart and S. S. Wilks, Editors. *1958. 268 pages. $6.00.*

READINGS in LINEAR PROGRAMMING

By S. VADJA, *Royal Naval Scientific Service*. Presents practical examples from the literature, as well as problems representative of actual practices. *1958. Approx. 100 pages. Prob. $3.00.*

Send for your examination copies.

JOHN WILEY & SONS, Inc. 440 Fourth Avenue, New York 16, N.Y.

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

VOLUME 53 JUNE, 1958 NUMBER 282

ARTICLES

Measuring Recessions . . . . . GEOFFREY H. MOORE 259
The Accuracy and Structure of Industry Expectations in Relation to Those of Individual Firms . . . . . ROBERT FERBER 317
An Index of Manufacturing Production in New England . . . . . HARRY BENJAMIN ERNST 336
Effect of Varying Degrees of Transitory Income on Income Elasticity of Expenditures . . . . . MARILYN DUNSING AND MARGARET G. REID 348
The First 1,945 British Steamships . . J. R. T. HUGHES AND STANLEY REITER 360
The Use of Random Work Sampling for Cost Analysis and Control . . . . . A. C. ROSANDER, H. E. GUTERMAN, AND A. J. MCKEON 382
The Trentile Deviation Method of Weather Forecast Evaluation . . . . . MORRIS JAMES SLONIM 398
Weight-Height Standards Based on World War II Experience . . . . . BERNARD D. KARPINOS 408
Actuarial Estimation of Survivorship in Chronic Disease . . . LILA ELVEBACK 420
Graphic Computation of Tau as a Coefficient of Disarray . . HAROLD D. GRIFFIN 441
A New Bivariate Sign Test . . . . . ISADORE BLUMEN 448
Nonparametric Estimation from Incomplete Observations . . . . . E. L. KAPLAN AND PAUL MEIER 457
Inadmissible Samples and Confidence Limits . . . . . HOWARD L. JONES 482
The Precision of Unbiased Ratio-Type Estimators . . . . . LEO A. GOODMAN AND H. O. HARTLEY 491
On Noncoverage of Sample Dwellings . . . . LESLIE KISH AND IRENE HESS 509
Complete Counterbalancing of Immediate Sequential Effects in a Latin Square Design . . . . . JAMES V. BRADLEY 525
Factorial Experimentation in Scheffé's Analysis of Variance for Paired Comparisons . . . . . OTTO DYKSTRA, JR. 529
Design and Operation of a Double-Limit Variables Sampling Plan . . . . . ACHESON J. DUNCAN 543
Empiric Investigation of a Test of Homogeneity for Populations Composed of Normal Distributions . . . . . G. A. BAKER 551
NOTES ABOUT AUTHORS. 558
SUMMARIES OF PAPERS DELIVERED AT 117th ANNUAL MEETING . 565
PUBLICATIONS RECEIVED 628
REPORT OF THE BOARD OF DIRECTORS AND SECRETARY-TREASURER FOR 1957 629

BOOK REVIEWS

TATE, MERLE W., AND CLELLAND, RICHARD C., *Nonparametric and Shortcut Statistics* . . . . . WILLIAM H. KRUSKAL 595
HENDRICKS, WALTER A., *The Mathematical Theory of Sampling* . T. A. BANCROFT 598
WIESEN, J. M., *Extension of Existing Tables of Attributes Sampling Plans* . . . . . DAVID R. HOWES 599
OWEN, D. B., *Table of the Normal Probability Integral* . . . HANS O. BREMER 600
METZGER, WILLIAM T., *Table of 10,000 Square Roots; Table of 45,000 Per Cents* . 600
SIMON, HERBERT A., *Models of Man* . . . . . MAY BRODBECK 600
BLANK, DAVID M., AND STIGLER, GEORGE J., *The Demand and Supply of Scientific Personnel* . . . . . M. BRONFENBRENNER 603

(continued on page iv)

SPANGLER, MILLER B., *New Technology and the Supply of Petroleum* . . . . . JOHN S. McGEE 604

TROY, LEO, *Distribution of Union Membership among the States, 1939 and 1953*. . . . . MARGARET S. GORDON 606

RATCLIFF, RICHARD U., RATHBUN, DANIEL B., AND HONNOLD, JUNIA, *Residential Finance, 1950* . . . . . SAUL B. KLAMAN 606

BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM, *Consumer Instalment Credit*, Part II, Vols. 1 and 2 . . . . . VIOLA WYCKOFF 609

MACK, RUTH P., *Consumption and Business Fluctuations* . . . DANIEL B. SUITS 610

MOSER, C. A., *The Measurement of Levels of Living with Special Reference to Jamaica* . . . . . ALICE HANSON JONES 611

THE MARKET RESEARCH SOCIETY, *Statistical Sources for Market Research* . . . . . PETER B. KENEN 613

KENDALL, MAURICE G., Editor, *The Sources and Nature of the Statistics of the United Kingdom*, Volume 2 . . . . . PETER B. KENEN 613

FUCHS, VICTOR R., *The Economics of the Fur Industry* . . . . G. R. GREGORY 614

KENDRICK, JOHN W., Editor, *Problems in the International Comparison of Economic Accounts* . . . . . EDGAR O. EDWARDS 615

SPENGLER, JOSEPH J., AND DUNCAN, OTIS DUDLEY, Editors, *Demographic Analysis: Selected Readings* . . . . . GEORGE J. STOLNITZ 617

MINISTRY FOR SOCIAL AFFAIRS, SUDAN GOVERNMENT, *The 1953 Pilot Population Census for the First Population Census in Sudan* . . W. PARKER MAULDIN 618

BRILLOUIN, LEON, CHERRY, COLIN, AND ELIAS, PETER, Editors, *Information and Control*, Volume 1, No. 1 620

MCCLOSKEY, JOSEPH F., AND COPPINGER, JOHN M., Editors, *Operations Research for Management*, Volume II . . . . . W. W. COOPER 621

MORRIS, WILLIAM T., *The Rationalization of Industrial Decision Processes* . . . . . CHARLES C. HOLT 624

HAMMER, PRESTON C., Editor, *The Computing Laboratory in the University* . . . . . FRANZ L. ALT 624

BARNES, RALPH M., *Work Sampling*, Second Edition . . . W. GRANT IRESON 625

BLANC-LAPIERRE, A., AND FORTET, ROBERT, *Théorie des Fonctions Aléatoires* . . . . . LEO J. TICK 626

1958 OFFICERS, AMERICAN STATISTICAL ASSOCIATION

Board of Directors

President	*President-Elect*	*Past President*
WALTER E. HOADLEY, JR.	RENSIS LIKERT	WILLIAM R. LEONARD

Vice-Presidents

MARTIN R. GAINSBRUGH	CHURCHILL EISENHART	HOWARD L. JONES

Directors

LESTER R. FRANKEL	H. O. HARTLEY	RALPH E. BURGESS
W. J. YOUDEN	GEOFFREY MOORE	OSCAR KEMPTHORNE

Secretary-Treasurer

DONALD C. RILEY

Members of the Council

HUGH H. BROWN	FRED C. LEONE	CHARLES SEDGEWICK
JEROME CORNFIELD	ARTHUR S. LITTELL	WILLIAM H. SHAW
CUTHBERT DANIEL	HERBERT MARKS	WALT R. SIMMONS
R. J. EGGERT	PHILIP J. MCCARTHY	JOHN H. SMITH
A. L. FINKNER	JOHN C. MCKEE	ALBERT T. SOMMERS
S. M. FREE	WESLEY D. MITCHELL	ARTHUR STEIN
IRWIN FRIEND	HORACE W. NORTON	HENRY TUCKER
BERNARD G. GREENBERG	ROY L. REIERSON	LEDYARD R. TUCKER
MORRIS HAMBURG	PHILIP J. RULON	W. ALLEN WALLIS
BOYD HARSHBARGER	JACQUES ST. PIERRE	EMMETT WELCH
ALLYN W. KIMBALL	MARVIN A. SCHNEIDERMAN	ALBERT W. WORTHAM
DUDLEY KIRK	MILTON F. SEARL	MARVIN ZELEN

Section Chairmen

HENRY TUCKER, Biometric Section
ROY L. REIERSON, Business and Economic Statistics Section
IRVING BURR, Section on Physical and Engineering Sciences
DUDLEY KIRK, Social Statistics Section
BOYD HARSHBARGER, Section on Training

Tables and other aids to computation appearing in this Journal are abstracted and indexed in *Mathematical Tables and Other Aids to Computation.*
Reprints of most articles published from 1945–1955 may be purchased from the Office of the Secretary.

EDITORIAL COLLABORATORS

FORMAN S. ACTON, *Princeton University*
OM P. AGGARWAL, *University of Saskatchewan*
STEPHEN G. ALLEN, *Stanford Research Institute*
R. L. ANDERSON, *University of North Carolina*
T. W. ANDERSON, *Columbia University*
KENNETH J. ARROW, *Stanford University*
EDWARD W. BARANKIN, *University of California (Berkeley)*
ROBERT BECHHOFER, *Cornell University*
AGNES BERGER, *Columbia University*
JOSEPH BERKSON, *Mayo Clinic*
JAMES N. BOLES, *University of California (Berkeley)*
GEORGE H. BORTS, *Brown University*
G. E. P. BOX, *Princeton University*
JOHN V. BREAKWELL, *North American Aviation Company*
IRWIN BROSS, *Cornell University Medical College*
MAX E. BRUNK, *Cornell University*
HOWARD BRUNSMAN, *Bureau of the Census*
JOSEPH BRYAN, *Massachusetts Institute of Technology*
MARION BRYSON, *Iowa State College*
B. H. CAMP, *Wesleyan University*
A. C. COHEN, JR., *University of Georgia*
W. S. CONNOR, *National Bureau of Standards*
WILLIAM A. CROMARTY, *Michigan State University*
LEE J. CRONBACH, *University of Illinois*
LEE CRUMP, *University of Rochester*
JOSEPH F. DALY, *Bureau of the Census*
HERBERT T. DAVID, *Iowa State College*
CYRUS DERMAN, *Columbia University*
W. J. DIXON, *University of California (Los Angeles)*
THOMAS G. DONNELLY, *Dominion Bureau of Statistics*
R. DOORNBOS, *Lever's Zeep-Maatschappij N. V. (Netherlands)*
HAROLD F. DORN, *National Institutes of Health*
ACHESON J. DUNCAN, *Johns Hopkins University*
DAVID B. DUNCAN, *University of North Carolina*
PAUL S. DWYER, *University of Michigan*
BENJAMIN EPSTEIN, *Stanford University*
WALTER T. FEDERER, *Cornell University*
JOHN FIRESTONE, *City College of New York*
LESTER R. FRANKEL, *Audits and Surveys Company*
LEO A. GOODMAN, *University of Chicago*
ROBERT A. GORDON, *University of California (Berkeley)*
BERT L. GREEN, *Massachusetts Institute of Technology*
B. G. GREENBERG, *University of North Carolina*
FRANK E. GRUBBS, *Ballistic Research Laboratories*
JOHN GURLAND, *Iowa State College*
MORRIS H. HANSEN, *Bureau of the Census*
BOYD HARSHBARGER, *Virginia Polytechnic Institute*
ALBERT GAILORD HART, *Columbia University*
H. O. HARTLEY, *Iowa State College*
DALE E. HATHAWAY, *Michigan State University*
W. BRADDOCK HICKMAN, *American Airlines*
J. L. HODGES, JR., *University of California (Berkeley)*
PAUL HORST, *University of Washington*
HAROLD HOTELLING, *University of North Carolina*
H. S. HOUTHAKKER, *Massachusetts Institute of Technology*
THOR HULTGREN, *National Bureau of Economic Research*
J. S. HUNTER, *American Cyanamid Company*
LOUIS C. HUNTER, *Industrial College of the Armed Forces*
LEONID HURWICZ, *University of Minnesota*
G. S. JAMES, *University of Leeds*
GEORGE JASZI, *Brookings Institution*
F. THOMAS JUSTER, *Amherst College*
A. W. KIMBALL, *Oak Ridge National Laboratory*
BRADFORD F. KIMBALL, *Public Service Commission (State of New York)*
HARRY V. KINCAID, *Stanford Research Institute*
WILLIAM H. KRUSKAL, *University of Chicago*
ANDRÉ C. LAURENT, *Wayne State University*
ERICH L. LEHMANN, *University of California (Berkeley)*
GERALD J. LIEBERMAN, *Stanford University*
FREDERIC M. LORD, *Educational Testing Service*
RUTH P. MACK, *National Bureau of Economic Research*
ALBERT MADANSKY, *Rand Corporation*
FRANK J. MASSEY, JR., *University of California (Los Angeles)*
QUINN MCNEMAR, *Stanford University*
PAUL E. MEEHL, *University of Minnesota*
PAUL MEIER, *University of Chicago*
MORRIS MENDELSON, *National Bureau of Economic Research*
SAMUEL MESSICK, *Educational Testing Service*
JAMES N. MORGAN, *University of Michigan*
HAROLD NISSELSON, *Bureau of the Census*
J. A. NORDIN, *Iowa State College*
DOUGLASS C. NORTH, *University of Washington*
EDWIN G. OLDS, *Carnegie Institute of Technology*
PAUL S. OLMSTEAD, *Bell Telephone Laboratories*
GUY H. ORCUTT, *Harvard University*
E. S. PEARSON, *University of London*
FRANK PROSCHAN, *Electronic Defense Laboratory*
HOWARD RAIFFA, *Harvard University*

GEORGE J. RESNIKOFF, *Illinois Institute of Technology*
HARRY M. ROSENBLATT, *Office of Naval Research*
JOAN R. ROSENBLATT, *National Bureau of Standards*
S. N. ROY, *University of North Carolina*
L. J. SAVAGE, *University of Chicago*
HENRY SCHEFFÉ, *University of California (Berkeley)*
CALVIN P. SCHMID, *University of Washington*
JACOB S. SIEGEL, *Bureau of the Census*
MONROE G. SIRKEN, *National Office of Vital Statistics*
TYNAN SMITH, *Federal Reserve System*
MILTON SOBEL, *Bell Telephone Laboratories*
STEPHEN SOBOTKA, *Northwestern University*
T. M. STANBACK, JR., *New York University*
JOSEPH STEINBERG, *Bureau of the Census*
FREDERICK F. STEPHAN, *Princeton University*
DANIEL B. SUITS, *University of Michigan*
B. V. SUKHATME, *Indian Council of Agricultural Research (New Delhi)*
ROBERT SUMMERS, *Yale University*
D. TEICHROEW, *Stanford University*
JAMES TOBIN, *Yale University*
MARY N. TORREY, *Bell Telephone Laboratories*
DONALD R. TRUAX, *University of Kansas*
M. C. K. TWEEDIE, *University of Manchester (England)*
HELEN M. WALKER, *Columbia University*
DAVID L. WALLACE, *University of Chicago*
VINCENT I. WEST, *University of Illinois*
D. RANSOM WHITNEY, *Ohio State University*
EVAN J. WILLIAMS, *North Carolina State College*
MARVIN ZELEN, *National Bureau of Standards*

Now Ready!

The 1957 Proceedings of the Business & Economic Statistics Section of the American Statistical Association

. . . includes 22 sessions covering topics dealing with many different phases of the Nation's economy. The papers were presented at the 117th Annual Meeting of the American Statistical Association held in Atlantic City, New Jersey, September 1957. The volume has 469 pages and is 8½" x 11" in size. The sessions, each of which includes papers and discussions by outstanding economists and statisticians, are as follows:

- How to Communicate Statistical Findings to Management
- Developments in Regional Analysis
- Stock Market Forecast for 1957-58
- Problems of Seasonal Measurement
- Factors Influencing Inventory Changes
- Consumer Expenditure Estimates for Food and Their Uses
- Income Distribution and Sub-Standard Levels of Living
- Some New Factors in the Long-Run Economic Outlook
- Developments in Methods for Short-Term Forecasting—I: A Review of National Bureau of Economic Research Lead Indicators and Other Diffusion Indexes
- Developments in Methods for Short-Term Forecasting — II: Correlation Techniques and Anticipatory Data
- Developments in Methods for Short-Term Forecasting—III: Econometric and Other Techniques
- The Economic Outlook
- Concentration in Manufacturing Industry
- Methods of Estimating Retail and Consumer Inventories
- Revision of Industrial Classification
- National Long-Term Economic Projections—I: Labor Force and Rate of Utilization
- National Long-Term Economic Projections—II: Developments in Projecting Productivity Trends
- Tolerances in Economic Projections and Forecasts
- Capital Formation
- Analysis of Consumption: Comparison of Aggregative Methods with Cross-Sectional Analysis of Survey Data
- Short-Term Demand for Consumer Durables and Housing: Research Methods
- Current Life Insurance Research on Owners and Buyers
- Contributed Papers

Price: to members, $3.25; to non-members, $3.75, *if remittance is sent with order.* An additional charge of $.25 will be made for postage and handling on orders received without remittance.

ORDER FORM

To: AMERICAN STATISTICAL ASSOCIATION, 1757 K St., N.W., Washington 6, D. C.

Please send me ..... copies of the 1957 Proceedings of The Business and Economic Statistics Section, @ $3.25 per copy for members; $3.75 per copy for non-members (if order is sent without remittance, $.25 will be added to the invoice to cover postage and handling).

☐ Payment Enclosed. ☐ Bill Company. ☐ Bill Me.

NAME _____

COMPANY _____

ADDRESS _____

CITY _____ ZONE _____ STATE _____

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 282 JUNE 1958 *Volume 53*

MEASURING RECESSIONS*

GEOFFREY H. MOORE
National Bureau of Economic Research

1. INTRODUCTION AND SUMMARY

THE purpose of this report is to provide a set of measurements of past business cycle recessions with which any current recession can be compared. The contractions in business activity that the American economy has experienced from time to time have, of course, varied widely in severity. Yet even severe depressions have often begun gradually. How soon can a severe decline be detected? How do the relative declines in the various available measures of economic activity compare with one another as a contraction develops? How can one determine whether a contraction that is currently under way is already or is going to be smaller or larger than those that have occurred in the past? How can one judge when it is about to end? The measurements presented here suggest possible ways of providing answers to questions such as these. Although the body of the report deals with measurements for recessions that have already run their full course, we shall, at the end, show how the method has worked out from month to month during the recession that began in 1957.

First, let us glance at the historical record of twenty-four business contractions given in Tables 260 and 261. The peak dates are the months when ex-

* Research along the lines developed in this report began during the recession of 1953–54, when tables comparing the percentage changes in a long list of economic series during the current and preceding recessions were prepared for Arthur F. Burns, then Chairman of the Council of Economic Advisers. In October 1957 R. J. Saulnier, present Chairman of the Council, requested the assistance of the National Bureau in preparing a similar set of tables. Tables covering some seventy monthly and quarterly economic series were promptly prepared. The Council compensated the National Bureau for the costs of this part of the statistical work. The electronic computer program was developed with the aid of a grant from the National Science Foundation. The International Business Machines Corporation generously contributed machine time on the 704 computer. These resources were essential to the pursuit of the study. No less essential were the intellectual and financial resources that have over the years been invested in the National Bureau's studies in business cycles. The present report is, in the truest sense, a product of these contributions, for the data and methods used here are virtually all derived in one way or another from this earlier work.

The statistical tables were prepared by Sophie Sakowitz, Dorothy O'Brien, and Sandra Renaud. Charlotte Boschan prepared and tested the electronic computer program. Alexander Pitts developed the materials necessary for the selection of the 1957 business cycle peak date and reviewed all the other peak dates. The charts were drawn by H. Irving Forman. Mary Phelps edited the manuscript. I am greatly indebted to these individuals as well as to others of the National Bureau staff for their wholehearted efforts to complete the job as speedily as possible. The comments of many who read the manuscript aided in its revision. These included Leo Grebler, Philip Klein, Maurice W. Lee, Ruth P. Mack, Roland I. Robinson, W. Allen Wallis, Donald B. Woodward, and Victor Zarnowitz. I am especially obliged to Arthur F. Burns, Solomon Fabricant, and Julius Shiskin for their sound advice and stimulating comments at all stages of the work.

The paper has been approved for publication as a report of the National Bureau of Economic Research by the director of research and the Board of Directors of the National Bureau, in accordance with the resolution of the board governing National Bureau reports (see the *Annual Report* of the National Bureau of Economic Research). It is to be reprinted as No. 61 in the National Bureau's series of Occasional Papers.

TABLE 260

THE DURATION OF BUSINESS CYCLE EXPANSIONS AND CONTRACTIONS IN THE UNITED STATES, 1854–1957

Business Cycle			Duration of	
			Expansion	Contraction
Trough	Peak	Trough	(months)	
Dec. 1854	June 1857	Dec. 1858	30	18
Dec. 1858	Oct. 1860	June 1861	22	8
June 1861	Apr. 1865	Dec. 1867	46	32
Dec. 1867	June 1869	Dec. 1870	18	18
Dec. 1870	Oct. 1873	Mar. 1879	34	65
Mar. 1879	Mar. 1882	May 1885	36	38
May 1885	Mar. 1887	Apr. 1888	22	13
Apr. 1888	July 1890	May 1891	27	10
May 1891	Jan. 1893	June 1894	20	17
June 1894	Dec. 1895	June 1897	18	18
June 1897	June 1899	Dec. 1900	24	18
Dec. 1900	Sep. 1902	Aug. 1904	21	23
Aug. 1904	May 1907	June 1908	33	13
June 1908	Jan. 1910	Jan. 1912	19	24
Jan. 1912	Jan. 1913	Dec. 1914	12	23
Dec. 1914	Aug. 1918	Mar. 1919*	44	7
Mar. 1919*	Jan. 1920	July 1921*	10	18
July 1921*	May 1923	July 1924	22	14
July 1924	Oct. 1926	Nov. 1927*	27	13
Nov. 1927*	Aug. 1929*	Mar. 1933	21	43
Mar. 1933	May 1937	June 1938*	50	13
June 1938*	Feb. 1945	Oct. 1945	80	8
Oct. 1945	Nov. 1948	Oct. 1949	37	11
Oct. 1949	July 1953	Aug. 1954	45	13
Aug. 1954	July 1957		35	
Average, 24 cycles, 1854–1954			29.9	19.9

For an explanation of the method used to determine the business cycle peak and trough dates and some tests of their validity, see Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (National Bureau of Economic Research, 1946), Ch. 4. A few of these dates (designated by an asterisk) have been revised since the Burns-Mitchell report, and the list has been carried forward to date.

* Revised.

pansion of aggregate economic activity culminated and contraction began, as judged from a variety of statistical records; the trough dates specify when contraction culminated and expansion began. The measures of duration show that five of the eight contractions since 1920 have lasted roughly a year (11 to 14 months). One was somewhat shorter (8 months), one somewhat longer (18 months), and one very much longer (43 months). Before 1920, contractions frequently lasted more than a year—indeed, ten out of sixteen between 1857 and 1919 lasted 18 months or more. The reasons for the apparent reduction in the typical length of contraction are not fully known, and we do not know whether it can be counted on as a permanent shift.[1] The intervening intervals

[1] For some observations suggesting that the durations of business contractions (and expansions) may be subject to long swings associated with the construction cycle and related developments, see the report by Moses Abramovitz in the National Bureau's *Thirty-eighth Annual Report* (May 1958).

TABLE 261

MEASURES OF THE DURATION AND AMPLITUDE OF BUSINESS CYCLE CONTRACTIONS IN THE UNITED STATES, 1920–1954

Business Cycle Peak	Business Cycle Trough	Mos. from Peak to Trough	Unemployment Rate, Annual: At Peak[a] (per cent)	Unemployment Rate, Annual: At Trough[a] (per cent)	Unemployment Rate, Annual: Change in Rate[a]	Unemployment Rate, Monthly: At Peak[b] (per cent)	Unemployment Rate, Monthly: At Trough[b] (per cent)	Unemployment Rate, Monthly: Change in Rate[b]	Nonagricultural Employment (percentage change from peak to trough[b])	Industrial Production (percentage change from peak to trough[b])	Gross National Product (percentage change from peak to trough[b])	Personal Income (percentage change from peak to trough[b])	Bank Debits Outside N.Y.C. (percentage change from peak to trough[b])	Retail Sales (percentage change from peak to trough[b])
1. Contractions in Chronological Order														
Jan. 1920	July 1921	18	4.0	11.9	+7.9	—	—	—	—	−29.0	—	—	−22.5	−4.2
May 1923	July 1924	14	3.2	5.5	+2.3	—	—	—	—	−16.3	−2.3[c]	+0.1[c]	−3.1	−1.7
Oct. 1926	Nov. 1927	13	1.9	4.1	+2.2	—	—	—	—	−5.7	+0.3[c]	+0.8[c]	+8.7	0
Aug. 1929	Mar. 1933	43	3.2	23.6	+20.4	0.58	24.96	+24.4	−30.7	−50.1	−49.6[c]	−49.8	−61.9	−43.3
May 1937	June 1938	13	14.3	19.0	+4.7	11.53	19.77	+8.2	−10.0	−31.5	−11.9[c]	−11.2	−16.5	−14.1
Feb. 1945	Oct. 1945	8	1.2	3.9	+2.7	0.11	0.34	+0.2	−7.9	−29.4	−10.9[c]	−4.1	−1.0	+8.7
Nov. 1948	Oct. 1949	11	3.4	5.5	+2.1	3.96	6.99	+3.0	−4.1	−7.7	−3.2[c]	−3.7	−5.3	−0.3
July 1953	Aug. 1954	13	2.5	5.0	+2.5	2.63	5.93	+3.3	−3.4	−9.5	−2.0[c]	−0.2	+0.2	−0.8
July 1957						4.23								
2. Contractions in Order of Severity (excl. February–October 1945)														
Oct. 1926	Nov. 1927	13	1.9	4.1	+2.2	—	—	—	—	−5.7	+0.3	+0.8	+8.7	0
July 1953	Aug. 1954	13	2.5	5.0	+2.5	2.63	5.93	+3.3	−3.4	−9.5	−2.0	−0.2	+0.2	−0.8
Nov. 1948	Oct. 1949	11	3.4	5.5	+2.1	3.96	6.99	+3.0	−4.1	−7.7	−3.2	−3.7	−5.3	−0.3
May 1923	July 1924	14	3.2	5.5	+2.3	—	—	—	—	−16.3	−2.3	+0.1	−3.1	−1.7
Jan. 1920	July 1921	18	4.0	11.9	+7.9	—	—	—	—	−29.0	—	—	−22.5	−4.2
May 1937	June 1938	13	14.3	19.0	+4.7	11.53	19.77	+8.2	−10.0	−31.5	−11.9	−11.2	−16.5	−14.1
Aug. 1929	Mar. 1933	43	3.2	23.6	+20.4	0.58	24.96	+24.4	−30.7	−50.1	−49.6	−49.8	−61.9	−43.3

All data are adjusted for seasonal variations. For sources, see Appendixes A and B. For a ranking of amplitudes of business cycle expansions and contractions from 1879 to 1933 see Table 156 in Burns and Mitchell, *op. cit.*

[a] Based on annual averages for the following business cycle peak and trough years: P, 1920; T, 1921; P, 1923; T, 1924; P, 1926; T, 1927; P, 1929; T, 1932; P, 1937; T, 1938; P, 1944; T, 1946; P, 1948; T, 1949; P, 1953; T, 1954; P, 1957. The annual data are from Stanley Lebergott, "Annual Estimates of Unemployment in the United States, 1900–1954," in *The Measurement and Behavior of Unemployment* (National Bureau of Economic Research, Special Conference Series 8, 1957), pp. 215–16.

[b] Based on three-month averages centered on business cycle peak and trough months, except as noted. The peak standings from which the percentage changes are computed are given in Appendix A.

[c] Based on changes between the following business cycle peak and trough quarters: P, I 1920; T, III 1921; P, II 1923; T, III 1924; P, III 1926; T, IV 1927; P, III 1929; T, I 1933; P, II 1937; T, II 1938; P, I 1945; T, IV 1945; P, IV 1948; T, IV 1949; P, II 1953; T, III 1954; P, III 1957.

of expansion have typically been substantially longer than the contractions, many of them lasting two or three years. This has been just as true since 1920 as it was before.

The durations of the expansions and the contractions are not sufficiently uniform or regular to give one more than a very rough notion about how long an expansion or contraction might be expected to last when it has just begun. After a year or so has elapsed, however, it may be of some help to know how frequently or infrequently phases of given lengths occur. Thus at the end of 1956 one could say this about the expansion that had begun in August 1954:

"If the current expansion were to continue through all of 1957, it will have lasted forty months. In the National Bureau's business cycle chronology covering the past 100 years there are only five expansions (out of twenty-four) that lasted as long as forty months: June 1861–April 1865, forty-six months; December 1914–August 1918, forty-four; March 1933–May 1937, fifty; June 1938–February 1945, eighty; and October 1949–July 1953, forty-five. Four of these expansions encompassed major wars, and one was the recovery from the Great Depression. Clearly, if the present expansion extends through 1957 without a setback it will establish a new precedent."[2] This bit of information in itself, of course, was not enough to forecast a recession, but it could usefully be considered together with other more direct and more important evidence. It now appears that the expansion came to an end in July 1957, that is, after 35 months (see below).

Table 261 shows the size of the declines between the business cycle peak and trough dates since 1920, as registered by several widely used measures of business activity. Clearly, a contraction that appears more severe than another by one measure may appear less severe by another measure. The 1953–54 contraction was somewhat greater than 1948–49 when measured in terms of the percentage decline in industrial production or the increase in the unemployment rate, but somewhat less than 1948–49 when measured by the percentage decline in gross national product or in nonagricultural employment.[3] Nevertheless, one can construct at least a rough ranking of the contractions according to severity.

At the top of the list of recent contractions, obviously, is the contraction that began in 1929—the longest and deepest. The only other that comes close to it in the National Bureau's 100-year chronology is the contraction of 1873 to 1879. Next most severe among those since 1920 are the contractions of 1920–21 and 1937–38. Both were very sharp and fairly short, but that of 1937–38 began when unemployment was still at a very high level, much higher than in 1920. The contraction of 1923–24 was of moderate amplitude, not unlike that of 1948–49. The most recent contraction, 1953–54, was in most respects of slightly lesser magnitude than the contraction of 1948–49, yet greater than that of 1926–27, and certainly more widely recognized. There remains the brief contraction after World War II, February–October 1945, which marked the transition from a wartime to a peacetime economy, and which is the most difficult

[2] *Thirty-seventh Annual Report*, National Bureau of Economic Research (May 1957), p. 53.

[3] Note that the table measures the percentage declines between business cycle peak and trough dates. Somewhat different results would be obtained if the declines were measured from the specific peak in each indicator to its trough. The latter method has some advantages for the purpose of measuring the amplitude of business cycles, and some disadvantages. In practice we have used both methods.

of all to characterize because different measures yield such different results. However, in terms of its impact upon the well-being of the population it must surely be classed among the more modest of those in our list.[4]

In order to have a definite scale we shall use the following ranking of contractions according to severity, excluding the 1945 episode because of its special character. The ranking is based partly on the information in Table 261 (see second section of table) and partly on other information bearing on the depth of these contractions (see note 3 above).

Contraction	Rank
Oct. 1926 Nov. 1927	1 (mildest)
July 1953 Aug. 1954	2
Nov. 1948 Oct. 1949	3
May 1923 July 1924	4
Jan. 1920 July 1921	5
May 1937 June 1938	6
Aug. 1929 Mar. 1933	7 (most severe)

We can then construct, for each of these contractions, measures that show by how much business activity declined from the peak as the contraction continued, and compare such measures with the above ranking. Such measures should tell us at about what stage—that is, how many months after the contraction began—the relative severity of each recession became evident, and how it manifested itself in different aspects of economic activity, such as production, employment, incomes, prices. Similar measures constructed during the course of a current contraction can then be used to appraise its severity and its scope compared with earlier contractions.

One of the prerequisites for such an analysis is that the current contraction is known or is believed to have begun, so that the date from which it starts, i.e., the peak of the business cycle, can be fixed. Of course, such a date may be selected tentatively, when a contraction is only suspected. If the assumption turns out to be an error, the error need not long persist. Experience suggests that the date of the peak can be determined with reasonable accuracy fairly soon after it occurs.[5]

Study of materials developed along these lines and presented later in the report suggests the following tentative conclusions:

---

[4] Another period that might be considered a cyclical contraction is 1951–52. Many sectors of the economy suffered setbacks at this time, after the rapid upsurge in 1950 when the Korean War began. Nevertheless, although the rate of growth of aggregate economic activity slackened perceptibly there was no appreciable over-all decline in output, income, or employment, and no rise in unemployment, since the defense industries kept expanding rapidly. Therefore we do not consider it to be a business cycle contraction. For an analytic description of this period see Bert G. Hickman, *The Korean War and United States Economic Activity, 1950–1952* (National Bureau of Economic Research, Occasional Paper 49, 1955). For the definition of business cycles followed in our work, and a discussion of its historical application, see Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (National Bureau of Economic Research, 1946), Chapters 1 and 4.

[5] For a description of some methods for accomplishing this see Arthur F. Burns, *Frontiers of Economic Knowledge* (National Bureau of Economic Research, 1954), especially pp. 107–34, 179–80, and Geoffrey H. Moore, *Statistical Indicators of Cyclical Revivals and Recessions* (National Bureau of Economic Research, Occasional Paper 31, 1950).

1. When a business recession begins, most broad indicators of aggregate economic activity (production, employment, income, trade) show relatively slight declines, and during the first six months of the recession the magnitude of the declines bears little relation to the ultimate severity or depth of the recession.

2. About six months after a recession begins, the percentage declines from the peak month to the current month in most economic aggregates are smaller in mild recessions than in severe recessions, and this ranking is maintained in succeeding months with little change.

3. When such comparisons are made for types of economic data that typically begin declining before a recession starts (for example, new orders, construction contracts, the average workweek, stock prices) the distinction between mild and severe recessions begins to appear as early as three or four months after the recession begins, and is also substantially maintained in succeeding months.

4. Although frequently both mild and sharp business contractions have ended within about a year, the recovery to the previous peak level has been accomplished much more quickly after mild contractions. Hence the period of depressed activity has been much longer when the contraction proceeded at a rapid rate.

5. While the above conclusions suggest that a rough ordering of recessions according to severity can be made within four to six months after the onset, they do not imply that either the ultimate depth or the duration of recessions can be reliably forecast by this means. Many factors not taken into account by the method, such as governmental measures taken to combat depression, have an important bearing on the severity and duration of business contractions. The method appears useful primarily in providing a yardstick against which a current decline in various aspects of economic activity can be gauged, and thereby facilitating a more accurate and enlightened appraisal of what has already taken place. This in itself might facilitate the development of appropriate counter-cyclical programs.

6. Measures of the strength of various counter-cyclical factors (for example, unemployment compensation payments, increased governmental expenditures, easier credit terms, lower taxes) at similar stages of recession might be developed on the same plan as described here, although it is not attempted in this study. Such measures might be of assistance in judging the prospects for further business contraction or for a resumption of economic expansion.

7. Several months before a recession comes to an end and an upturn in aggregate activity occurs, a progressive narrowing of the scope of contraction ordinarily becomes visible. Fewer activities continue to decline, more begin to rise. It appears first in series of the "leading" type. The more extensive and more sustained this reduction in the scope of the contraction is, the more likely that it marks the real end of recession rather than an abortive recovery. Information of this sort may help to identify an upturn in aggregate activity at about the time it occurs or shortly thereafter.

8. When the methods developed in this investigation are applied to the business contraction that began in July 1957, we find that:

   a. After eight months of contraction, i.e., through March 1958, most indicators have declined more than in the corresponding periods of the four milder contractions

since 1920 (1923–24, 1926–27, 1948–49, 1953–54) and less than in the three more severe contractions (1920–21, 1929–30, 1937–38).

b. The intermediate position of the 1957–58 contraction first became apparent in data for the leading indicators for November 1957, i.e., four months after the peak of July 1957. It was confirmed by most indicators of aggregate economic activity when data for February 1958 became available.

c. In contractions of the severity indicated for the 1957–58 contraction, it would be in line with previous experience if the level of economic activity generally remained below the previous peak level (July 1957) for a period ranging from a year rand a half to two and a half years.

d. One of the outstanding features of the first eight months of the 1957–58 contraction has been the relatively modest decline in personal income. The rise in consumers' prices has been less unusual, since increases occurred during the first eight months of four of the seven business contractions since 1920.

9. The tentative findings reported above need to be tested further.[6] The method could usefully be tested on declines that did not reach business cycle proportions. Comparisons based on a different method of dating downturns—e.g., dating the downturn from the peak in the specific series being compared—should be made, and other ways of measuring the severity of recessions should be explored. The empirical results should be examined in the light of the hypotheses that have been advanced to account for variations in the severity of business cycle contractions. Work along these lines will be facilitated now because electronic computer programs are available to handle the computations.

2. CHANGES IN AGGREGATE ECONOMIC ACTIVITY DURING THE FIRST YEAR OF RECESSION

Table 266 shows how a comparison of developments during the first year of mild and severe business contractions works out for one widely used economic indicator, the Federal Reserve index of industrial production. Percentage changes are computed from the peak standing (a three-month average that includes the business cycle peak month, the month preceding and the month following) to one month after the peak, two months after the peak, and so on up to twelve months after the peak. The table covers the seven business cycle contractions since 1920 (excepting the contraction that followed World War II). Note that the peak dates are not necessarily those at which the production index reached its peak, but rather when business activity at large did so. Usually the peak in the production index has not differed by more than a month or two from the business cycle peak.[7] Use of the business cycle peak enables us to examine a wide variety of series on a comparable basis (see below).

---

[6] Two important contributions have already been made. Julius Shiskin has constructed and analyzed an extensive set of measures of the scope, magnitude and rate of change in the separate industry components of various economic aggregates such as employment, production, and new orders. He has compared the current contraction with those beginning in 1953, 1948, 1937, and 1929 on a plan similar to that used here, and also on a plan that uses the "specific cycle peak" dates in each aggregate as the point from which to start the comparison. This work has been carried on at the Bureau of the Census for the Council of Economic Advisers.

Pao Lun Cheng, Michigan State University, in a paper on "Statistical Indicators and Cyclical Amplitudes," presented at the annual meeting of the Midwest Economics Association, Des Moines, Iowa, on April 19, 1958, explores the relations between the severity of business cycle contractions and the rates of change in indicators prior to and during the contractions. Part of this work is along lines very similar to those followed here, and yields similar conclusions. In addition, however, Dr. Cheng tests a number of interesting hypotheses that go well beyond our own work.

[7] At one of the peaks the difference was 5 months; at one, 2 months; at three, 1 month; and at two there was no difference. In the current recession the difference is somewhat greater than usual. The peak in the production index, according to revised figures published in March 1958, was reached in December 1956 or February 1957 (146

TABLE 266

PERCENTAGE CHANGES IN INDUSTRIAL PRODUCTION DURING SEVEN BUSINESS CYCLE CONTRACTIONS

Months after Peak	Business Cycle Peak						
	Oct. 1926	July 1953	Nov. 1948	May 1923	Jan. 1920	May 1937	Aug. 1929
	Standing at Peak (1947–49: 100)[a]						
	52.7	136.3	103.7	49.0	42.7	64.7	60.7
	Percentage Change from Peak						
1	−1.3	−0.2	−1.6	0	+3.1	−1.0	−1.1
2	−1.3	−2.4	−2.6	−2.0	+0.8	+0.5	−2.7
3	−1.3	−3.2	−4.5	−2.0	−3.9	−1.0	−7.7
4	−1.3	−5.4	−6.4	−4.1	−1.6	−4.1	−11.0
5	+0.6	−7.6	−7.4	−6.1	−1.6	−11.9	−11.0
6	−1.3	−8.3	−8.4	−6.1	−3.9	−19.6	−11.0
7	−1.3	−8.3	−8.4	−8.2	−3.9	−27.3	−12.6
8	−1.3	−9.8	−9.3	−6.1	−6.2	−28.9	−12.6
9	−3.2	−9.8	−7.4	−4.1	−8.6	−30.4	−15.9
10	−3.2	−8.3	−6.4	−6.1	−18.0	−30.4	−17.6
11	−5.1	−9.0	−9.3	−8.2	−22.7	−32.0	−20.9
12	−5.1	−9.8	−7.4	−12.2	−27.3	−33.5	−24.2
	Rank of Percentage Change[b]						
1	6	3	7	2	1	4	5
2	3	5	6	4	1	2	7
3	2	4	6	3	5	1	7
4	1	5	6	3	2	4	7
5	1	5	4	3	2	7	6
6	1	4	5	3	2	7	6
7	1	4	5	3	2	7	6
8	1	5	4	2	3	7	6
9	1	5	3	2	4	7	6
10	1	4	3	2	6	7	5
11	1	3	4	2	6	7	5
12	1	3	2	4	6	7	5

[a] Three-month average, centered on the peak month.

[b] A rank of 1 is given to the smallest decline (or largest rise) among the seven contractions, a rank of 7 to the largest decline, etc.

Source: Board of Governors of the Federal Reserve System. Index is adjusted for seasonal variations. The seven business cycle contractions are arrayed from left to right according to their over-all severity (see text).

Table 266 makes it clear that changes in the volume of industrial output in the first month or two of a business recession are typically slight. Declines of only 2 or 3 per cent are the rule. Moreover, the relative severity of the decline manifests itself only irregularly during the first six months or so after the con-

in both months), seven or five months before the business cycle peak in July 1957. Before revision, the index showed December 1956 one point higher than February. The revised index declines to 144 in April and May, rises to 145 in June, July and August, and then declines to date.

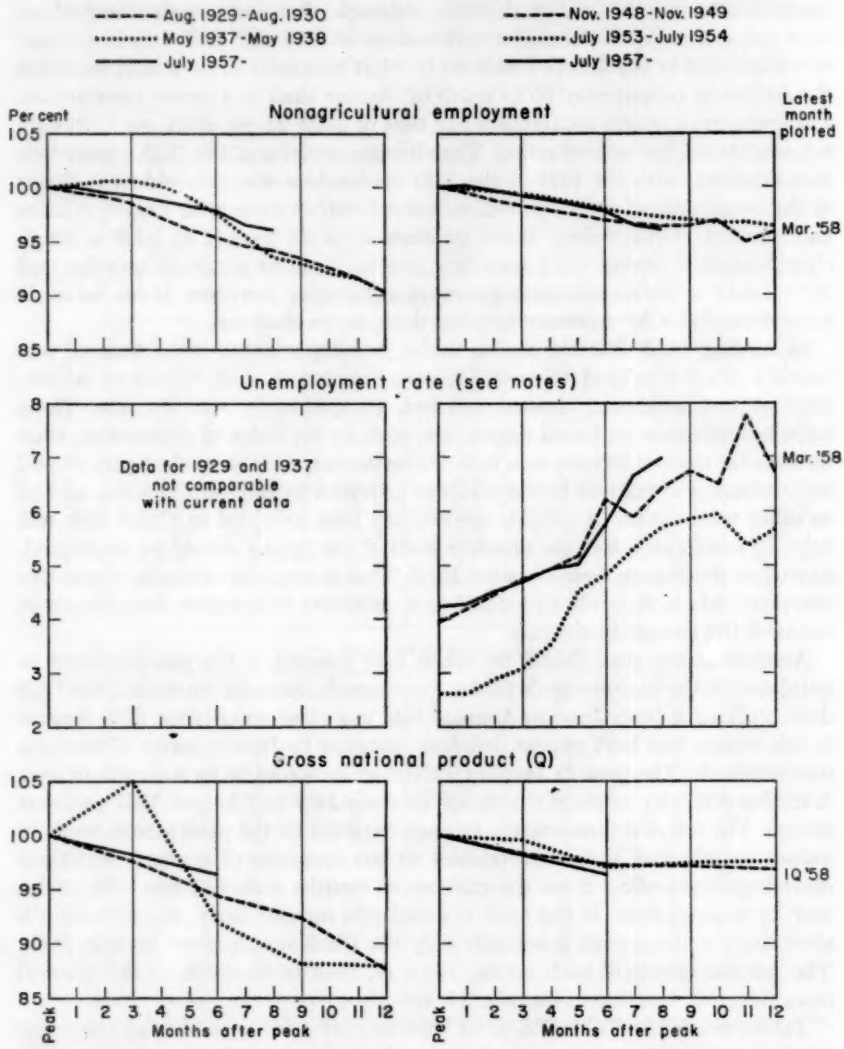
traction begins. The mild recessions of 1948 and 1953 began with a relatively sharp decline in industrial production, although after five months the declines were not as great as in the major contractions of 1929 and 1937. It is important to realize that in the first few months of what turns out to be a mild recession the decline in output may be as sharp or sharper than in a severe contraction. Furthermore, a severe contraction like that of 1920–21 can start out with only a moderate decline in production. The ultimate severity of the 1929 contraction in comparison with the 1921 or the 1937 contraction was not evident in terms of the magnitude of the decline in industrial output even after twelve months had elapsed. Nevertheless, these measurements do permit at least a rough classification of contractions according to severity after about six months, and the validity of the classification improves as the span increases. It can be made more dependable by reference to other data, as we shall see.

In reading Table 266 and similar tables it is important to recall that all economic series are subject to erratic movements due to such factors as strikes, international incidents, unusual weather, flu epidemics, and the like. These have less influence on broad aggregates, such as the index of production, than on data for narrow sectors, as a rule. Nevertheless, each monthly figure should be scrutinized in relation to the evidence provided by adjacent months, as well as other information. A graphic record, like that provided in Chart 268, will help. In some cases, too, the absolute level of the figures should be considered, as well as the change from the peak level. That is why, for example, the unemployment rate is shown in Chart 268 as an absolute rate, rather than simply in terms of the change in the rate.

Another factor that should be taken into account is the possible error in our dates for the business cycle peaks. For example, recently we revised the 1929 date, shifting it from June to August. July is a close competitor with August in this choice, but both appear definitely superior to June in terms of the data now available. The peak in January 1920 may be too early by a month or two. A similar difficulty exists in the choice between July and August 1957 (see text below). The use of a three-month average centered on the peak month tends to reduce the effect of such uncertainties on our measures of changes. But their more important effect is on the number of months a decline has been under way by a given date. If the peak is dated two months early, the fifth month after the erroneous peak is actually only the third month after the true peak. The relative effects of such errors, however, tend to diminish as the interval from the peak lengthens (see note 14, below).

Tables similar to Table 266 have been compiled for a number of other important aggregative measures of economic activity, and are reproduced in Appendix A. Here we shall consider only the rankings of the percentage changes (Table 273). Since these series figure significantly in our determination of business cycle turning dates, usually their peaks and that of the business cycle closely coincide. At any given turn, however, some may continue to rise for a few months, as the table indicates. On the other hand, some may have begun to decline before the designated peak month, so that in these cases the ranking is based on only that part of the decline that took place after the business cycle peak.

CHART 268. Relative changes after business cycle peaks, selected series.*



* All series except the unemployment rate are converted to indexes, using the three-month average centered on the business cycle peak months or quarters as the base (=100). The peak months (quarters) are: August (3Q) 1929, May (2Q) 1937, November (4Q) 1948, July (2Q) 1953, and July (3Q) 1957. All series except industrial stock prices and wholesale prices are adjusted for seasonal variations.

CHART 268 (*continued*)

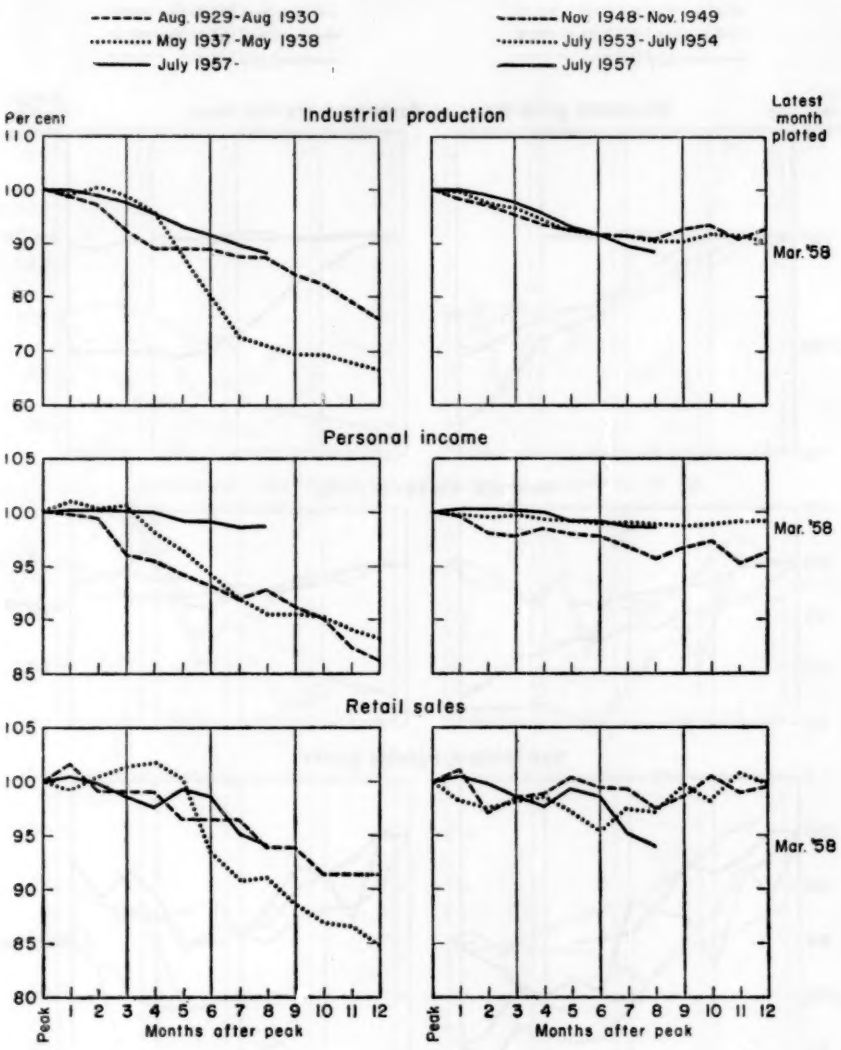


CHART 268 (*continued*)

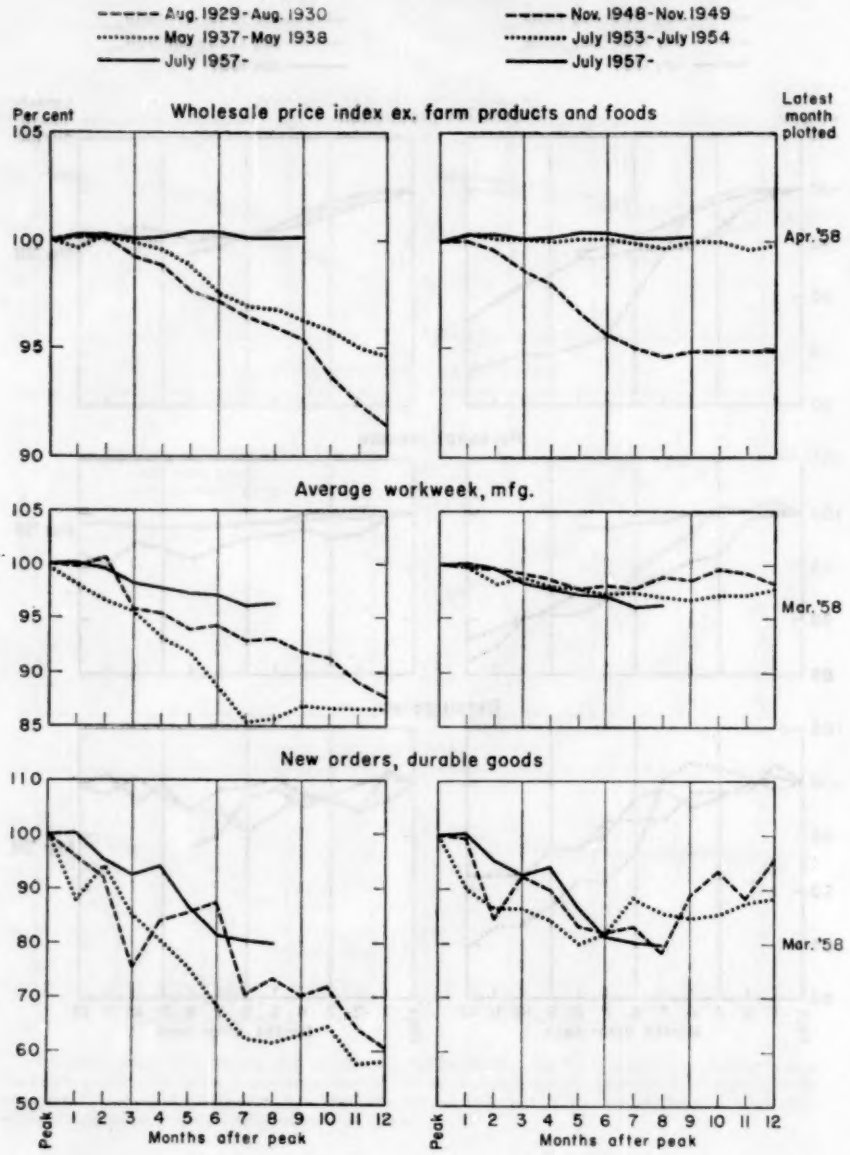
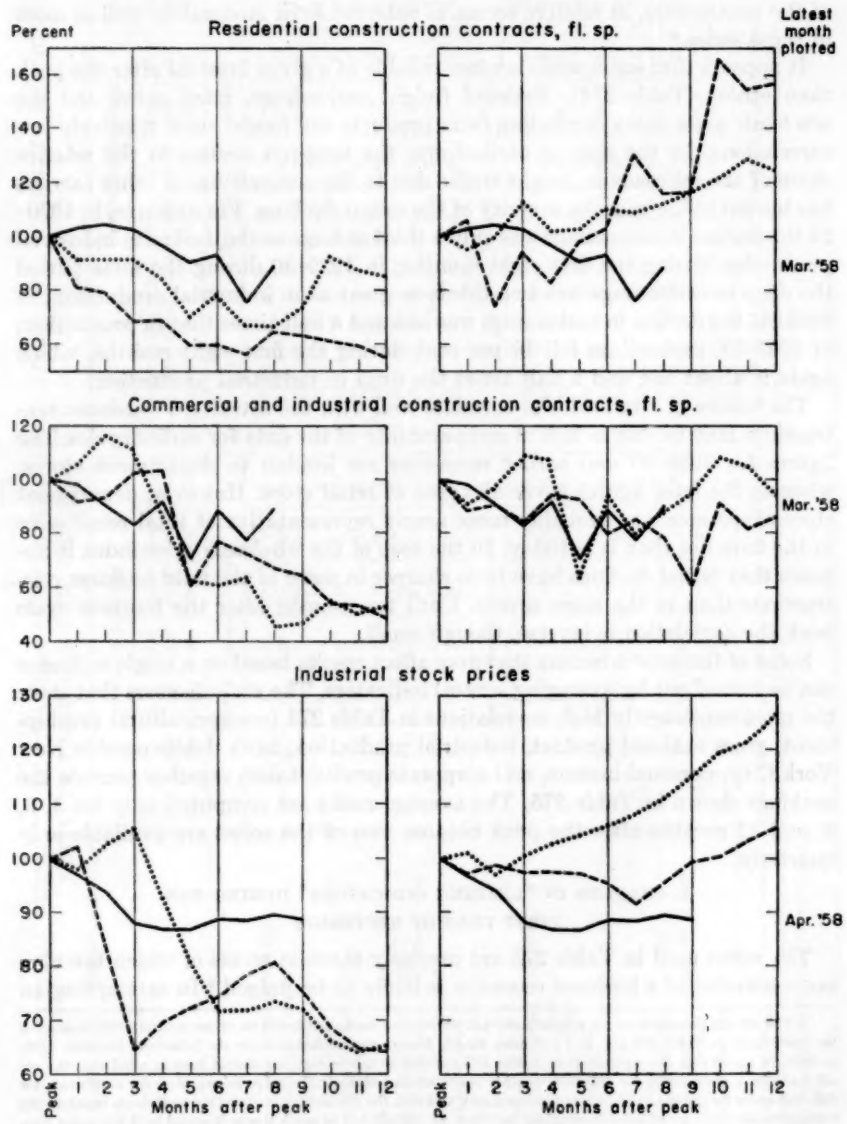


CHART 268 (*continued*)



On the whole, Table 273 confirms the showing of Table 266. The initial changes in measures of the general level of economic activity may be quite misleading if regarded as an indication of the severity or mildness of the impending decline. It is not until some six months after the peak that the ultimate severity of the contraction, in relative terms, is reflected even moderately well in most of these series.[8]

It appears that some series are less reliable at a given interval after the peak than others (Table 274). Railroad freight carloadings, retail sales, and the wholesale price index (excluding farm products and foods) show relatively low correlations. In the case of carloadings, the long-run decline in the relative share of the railroads in freight traffic due to the competition of other carriers has tended to increase the severity of the recent declines. For example, in 1920–21 the decline in carloadings was only a third as large as the decline in industrial production during the first eight months; in 1929–30 during the same period the drop in carloadings was two-thirds as great as in industrial production; in 1953–54 the decline in carloadings was one and a half times that in production; in 1957–58, carloadings fell 18 per cent during the first eight months, which again is about one and a half times the drop in industrial production.

The failure of retail sales to correlate well with the severity of business contractions may be due to lack of comparability of the data for earlier cycles. The figures for 1929–30 and earlier recessions are limited to department stores, whereas the later figures cover all types of retail store. However, department store experience was probably more nearly representative of total retail sales in the twenties than it is today. In the case of the wholesale price index it appears that initial declines have been sharper in some of the mild business contractions than in the more severe. Until ten months after the business cycle peak the correlation is inverse, though small.

Some of the erratic factors that may affect results based on a single indicator can be ironed out by averaging several indicators. The six indicators that show the most consistently high correlations in Table 274 (nonagricultural employment, gross national product, industrial production, bank debits outside New York City, personal income, and corporate profits) taken together provide the rankings shown in Table 275. The average ranks are computed only for 3, 6, 9, and 12 months after the peak because two of the series are available only quarterly.

3. CHANGES IN "LEADING INDICATORS" DURING THE FIRST YEAR OF RECESSION

The series used in Table 273 are precisely those in terms of which the ultimate severity of a business recession is likely to be judged.[9] In attempting an

[8] The .05 significance level for a Spearman rank correlation coefficient based on seven observations, as most of the coefficients in Table 273 are, is .71 (Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956, p. 284). In appraising the coefficients in Tables 273 through 276, however, one should keep in mind that the coefficients for a given series for different monthly spans are not statistically independent, that the coefficients for different series for the same span are not independent, and that the .05 significance level for coefficients based on the average ranks of groups of series is certainly less than .71, though not as much less as it would be if the series were independent. The coefficients are presented mainly to provide a convenient summary statistic on the degree of relationship between the changes in the indicators and the severity of the contractions.

[9] The number of unemployed or the unemployment rate should of course be considered in any such appraisal. They are omitted from Table 273 because the available monthly data for contractions before World War II are not comparable in magnitude with current data. The unemployment rate is included in Table 278, where only the directions of change are utilized.

TABLE 273

RANKING OF PERCENTAGE CHANGES IN SELECTED INDICATORS OF AGGREGATE ECONOMIC ACTIVITY DURING SEVEN BUSINESS CYCLE CONTRACTIONS

Indicator	Contraction Beginning							Rank Correlation with Severity of Contraction	Contraction Beginning							Rank Correlation with Severity of Contraction
	Oct. 1926 (1)	July 1953 (2)	Nov. 1948 (3)	May 1923 (4)	Jan. 1920 (5)	May 1937 (6)	Aug. 1929 (7)		Oct. 1926 (1)	July 1953 (2)	Nov. 1948 (3)	May 1923 (4)	Jan. 1920 (5)	May 1937 (6)	Aug. 1929 (7)	
	Rank of *3* month change								Rank of *6* month change							
1. Nonagricultural employment, BLS	–	2	4	–	–	1*	3	0	–	1	3	–	–	2	4	.80
2. Gross national product (Q)	2*	4	5	3*	–	1*	6	.26	1*	4	3	2	–	6	5	.71
3. Industrial production	2	4	6	3	5	1	7	.29	1	4	5	3	2	7	6	.64
4. Freight carloadings	1*	4	5	2	6	3	7	.61	2	6	4	1	3	7	5	.36
5. Bank debits outside N.Y.C.	3*	4	7	6	2*	5	1*	−.32	1*	4	5	3	2*	6	7	.68
6. Personal income	5	3	4	2	–	1*	6	.09	2	3	4	1*	–	5	6	.66
7. Retail sales	4	7	6	2*	1*	3*	5	−.36	4	6	2	3	1*	7	5	.14
8. Corporate profits after taxes (Q)	2	1	4	7	3	5	6	.68	1	3	2	5	4	7	6	.89
9. Wholesale price index, excl. farm & food	6	2*	5	7	1*	3	4	−.29	6	2*	5	7	1*	3	4	−.29
Average rank, 6 indicators[a]	2.8	3.0	5.0	4.2	3.3	2.3	4.8		1.2	3.2	3.7	2.8	2.7	5.5	5.7	
Rank of average rank	2	3	7	5	4	1	6	.18	1	4	5	3	2	6	7	.68
	Rank of *9* month change								Rank of *12* month change							
1. Nonagricultural employment, BLS	–	1	2	–	–	4	3	.80	–	1	2	–	–	4	3	.80
2. Gross national product (Q)	2	3	4	1*	–	6	5	.60	2*	3	4	1*	–	5	6	.66
3. Industrial production	1	5	3	2	4	7	6	.71	1	3	2	4	6	7	5	.86
4. Freight carloadings	3	5	6	2	1	7	4	.07	1	4	6	3	2	7	5	.50
5. Bank debits outside N.Y.C.	1*	3	5	2*	4	7	6	.79	1*	2	4	3	5	6	7	.96
6. Personal income	2*	3	4	1*	–	6	5	.60	1	2	4	3	–	5	6	.94
7. Retail sales	2.5	4	5	1*	2.5	7	6	.50	2	3	4	5	1*	7	6	.61
8. Corporate profits after taxes (Q)	4	1	3	2	7	6	5	.61	2	1	3	4	7	6	5	.82
9. Wholesale price index, excl. farm & food	7	2	6	4	1*	3	5	−.32	4	1	2	5	7	3	6	.54
Average rank, 6 indicators[a]	2.0	2.7	3.5	1.6	5.0	6.0	5.0		1.4	2.0	3.2	3.0	6.0	5.5	5.3	
Rank of average rank	2	3	4	1	5.5	7	5.5	.72	1	2	4	3	7	6	5	.82

Source: Appendix A. The seven business contractions are arrayed from left to right according to their over-all severity (see text). Changes in a given indicator during each of the seven contractions are ranked according to magnitude. The smallest decline (or largest rise) is assigned a rank of 1, the next smallest a rank of 2, etc. All series except the wholesale price index are adjusted for seasonal variations.

* Increase.

[a] Nonagricultural employment, gross national product, industrial production, bank debits, personal income, corporate profits.

TABLE 274

RANK CORRELATION COEFFICIENTS: SEVERITY OF BUSINESS CYCLE CONTRACTIONS AND PERCENTAGE CHANGES IN SELECTED INDICATORS

Indicator	No. of Contractions Covered[a]	Months after Peak											
		1	2	3	4	5	6	7	8	9	10	11	12
Indicators of Aggregate Economic Activity													
Nonagricultural employment, BLS	4	.40	0	0	.40	.40	.80	1.00	.80	.80	.60	.80	.80
Gross national product (Q)	6	—	—	.26	—	—	.71	—	—	.60	—	—	.66
Industrial production	7	−.25	.04	.29	.43	.61	.64	.64	.64	.71	.75	.79	.86
Freight carloadings	7	.54	−.68	.61	−.07	.29	.36	.07	.18	.07	.21	.21	.50
Bank debits outside N.Y.C.	7	.93	−.29	−.32	.54	.57	.68	.75	.75	.79	.75	.96	.96
Personal income	4,6[b]	−.60	0	.09	1.00	1.00	.66	1.00	.80	.60	1.00	1.00	.94
Retail sales	7	−.29	−.39	−.36	−.43	0	.14	.29	.29	.50	.50	.46	.61
Corporate profits after taxes (Q)	7	—	—	.68	—	—	.80	—	—	.61	—	—	.82
Wholesale price index, excl. farm & food	7	.04	−.32	−.29	−.29	−.29	−.29	−.29	−.43	−.32	.21	.43	.54
Leading Indicators													
Average workweek, mfg.	6	−.26	−.37	.43	.89	.89	.77	.89	.77	.77	.77	.77	.71
Layoff rate, mfg., *inverted*	7	−.04	−.07	.07	.04	.14	.14	.32	.29	.32	.43	.71	.71
Accession rate, mfg.	7	0	.46	.64	.46	.43	.39	.54	.21	.14	.25	.25	.25
New orders, durable goods	7	.21	−.14	.64	.71	.14	.21	.75	.68	.61	.61	.71	.75
Residential construction contracts	7	.79	.39	.82	.79	.64	.57	.57	.54	.54	.57	.64	.64
Comm. & indus. const. contracts	7	.04	.07	−.11	.21	.79	.82	.71	.82	.46	.79	.82	.86
New incorporations, number	7	.14	−.14	.61	.68	.75	.71	.82	.86	.68	.64	.71	.71
Business failures, liabilities, *inverted*	7	.39	.07	.54	.14	−.32	.14	.61	.25	.50	.61	.57	.50
Basic commodity price index	7	.07	.21	.14	.39	.57	.46	.46	.39	.68	.64	.64	.68
Industrial stock price index	7	0	.39	.39	.93	.96	.93	.89	.89	.89	.93	.96	.96

[a] See Tables 273 and 276.

[b] Entries for 3, 6, 9, and 12 months after peak include the two contractions for which quarterly data are available (1923–24 and 1926–27) as well as the four covered by monthly data.

TABLE 275

RANKING OF SEVEN BUSINESS CYCLE CONTRACTIONS ACCORDING TO SEVERITY IN SUCCESSIVE MONTHS AFTER PEAKS, TWO GROUPS OF INDICATORS

Months after Peak	Rank of Average Rank Contraction beginning							Rank Correlation with Severity of Contraction
	Oct. 1926 (1)	July 1953 (2)	Nov. 1948 (3)	May 1923 (4)	Jan. 1920 (5)	May 1937 (6)	Aug. 1929 (7)	
	Six Indicators of Aggregate Economic Activity[a]							
3	2	3	7	5	4	1	6	.18
6	1	4	5	3	2	6	7	.68
9	2	3	4	1	5.5	7	5.5	.72
12	1	2	4	3	7	6	5	.82
	Ten Leading Indicators[b]							
1	1	3.5	6	3.5	5	7	2	.32
2	1	2	7	3	6	5	4	.50
3	1	2	5	3.5	3.5	6	7	.88
4	1	2	3	4	6	5	7	.96
5	1	2	3	4	5	7	6	.96
6	1	3	4	2	5	7	6	.86
7	1	3	4	2	5	7	6	.86
8	1	2.5	4	2.5	6	7	5	.83
9	1	2.5	4	2.5	7	6	5	.79
10	2.5	2.5	1	4	7	6	5	.74
11	1	3	2	4	7	6	5	.82
12	3	1	2	4	7	5.5	5.5	.77

Source: Tables 273 and 276 and Appendix A. The seven business contractions are arrayed from left to right according to their over-all severity (see text). A rank of 1 indicates a small decline, a rank of 7 a large decline.

[a] Nonagricultural employment, gross national product, industrial production, bank debits, personal income, corporate profits (see text).

[b] See Table 276.

early estimate of relative severity, however, other types of series may be more helpful. In the National Bureau's studies of business cycles numerous series have been uncovered that typically turn down some months before the peak in general business. Because of their sensitivity, they may provide an early indication of the severity of contractions. Table 276 contains the rankings of the percentage changes for a group of them, based on the figures provided in Appendix A.

We find that even as early as the third month after the peak the declines in many of the leading series begin to differentiate the more severe from the less severe recessions. Hence something can be gained by watching the early movers. Again, some appear less reliable for the purpose than others. The layoff and accession rates, business failure liabilities, and new orders for durable goods show relatively low or erratic correlation coefficients (Table 274). Because of

TABLE 276

RANKING OF PERCENTAGE CHANGES IN TEN LEADING INDICATORS DURING SEVEN BUSINESS CYCLE CONTRACTIONS

Indicator	Contraction Beginning							Rank Correlation with Severity of Contraction	Contraction Beginning							Rank Correlation with Severity of Contraction
	Oct. 1926 (1)	July 1953 (2)	Nov. 1948 (3)	May 1923 (4)	Jan. 1920 (5)	May 1937 (6)	Aug. 1929 (7)		Oct. 1926 (1)	July 1953 (2)	Nov. 1948 (3)	May 1923 (4)	Jan. 1920 (5)	May 1937 (6)	Aug. 1929 (7)	
	Rank of *3* month change								Rank of *6* month change							
1. Average workweek, mfg.	4	3	1	2	—	6	5	.43	1	4	2	3	—	6	5	.77
2. Layoff rate, mfg., *inverted*	3	5	6	2	1*	7	4	.07	2	6	5	3	1	7	4	.14
3. Accession rate, mfg.	2	3	1	7	5	4	6	.64	2	4	3	7	1*	6	5	.39
4. New orders, durable goods	1*	5	4	3	2*	6	7	.64	2	6	5	1	3	7	4	.21
5. Residential construction contracts	2	1*	5	4	3	6	7	.82	4	2*	3*	1*	7	5	6	.57
6. Comm. & indus. const. contracts	4	2*	5	6	7	1*	3*	−.11	1*	2	4	3	7	6	5	.82
7. New incorporations, number	2	1*	6	3	5	7	4	.61	1*	4	3	2	7	6	5	.71
8. Business failures, liabilities, *inverted*	4	3	2	1*	7	5	6	.54	6	1*	2*	4	7	5	3	.14
9. Basic commodity price index	2	3	7	6	1*	5	4	.14	3	1*	7	2	4	6	5	.46
10. Industrial stock price index	3	2*	4	6	5	1*	7	.39	2*	1*	3	4	5	7	6	.93
Average rank, 10 indicators	2.7	2.8	4.1	4.0	4.0	4.8	5.3		2.4	3.1	3.7	3.0	4.7	6.1	4.8	
Rank of average rank	1	2	5	3.5	3.5	6	7	.88	1	3	4	2	5	7	6	.86
	Rank of *9* month change								Rank of *12* month change							
1. Average work week, mfg.	2	3	1	4	—	6	5	.77	3	2	1	4	—	6	5	.71
2. Layoff rate, mfg., *inverted*	1	6	4	2	5	7	3	.32	1	2	5	3	6	7	4	.71
3. Accession rate, mfg.	3	4	2	7	6	1	5	.14	4	3	1	6	7	2	5	.25
4. New orders, durable goods	2	4	3	1*	7	6	5	.61	2	3	1	4	7	6	5	.75
5. Residential construction contracts	4	3*	2*	1*	7	5	6	.54	4	2*	1*	3*	7	5	6	.64
6. Comm. & indus. const. contracts	3	2	5	1*	7	6	4	.46	2	1	3	4	7	5	6	.86
7. New incorporations, number	3	1*	2*	5	6	7	4	.68	3*	1*	2*	5	7	4	6	.71
8. Business failures, liabilities, *inverted*	4	3	2	1*	7	6	5	.50	4	3*	2*	1*	7	6	5	.50
9. Basic commodity price index	2	1*	5	3	6	7	4	.68	1*	2*	5	3	7	6	4	.68
10. Industrial stock price index	2*	1*	4	3*	5	7	6	.89	2*	1*	3*	4	5	6	7	.96
Average rank, 10 indicators	2.6	2.8	3.0	2.8	6.2	5.8	4.7		2.6	2.0	2.4	3.7	6.7	5.3	5.3	
Rank of average rank	1	2.5	4	2.5	7	6	5	.79	3	1	2	4	7	5.5	5.5	.77

Source: Appendix A. The seven business contractions are arrayed from left to right according to their over-all severity (see text). Changes in a given indicator during each of the seven contractions are ranked according to magnitude. The smallest decline (or largest rise) is assigned a rank of 1, the next smallest a rank of 2, etc. The layoff rate and business failures are treated invertedly, i.e., the smallest increase (or largest decline) gets a rank of 1, etc. All series except basic commodity prices and industrial stock prices are adjusted for seasonal variation.

* Increase (in layoff rate and business failures, decrease).

the erratic month-to-month changes that many of these series exhibit, it is especially desirable to study all the evidence rather than rely on data for a single month or a single series. The average ranks for all ten leading indicators provide one means for doing this (Table 275), though of course the individual series and their relations to one another and to other series should be studied as well. As a group, the leaders provide a more reliable indication of severity than the aggregates during the first six or possibly nine months of recession.[10]

4. CHANGES IN THE SCOPE OF BUSINESS CYCLE CONTRACTIONS

The measures provided in Appendix A can help, also, to show how widespread the economic contraction is and whether it is spreading further. Table 278 shows how this may be done, using ten indicators of aggregate economic activity and ten leading indicators. The left-hand section of the table indicates, for successive months after the business cycle peak, what proportion of the series were above their level at the cycle peak (that is, a three-month average centered on the peak date). Very few of the aggregates (section A) reached such a favorable position during the first twelve months of the several recessions, although in the milder recessions one or two of the aggregates could be counted in this group in nearly every month. In 1953–54 eighteen months elapsed before half the aggregates regained the levels they had attained at the business cycle peak, and in 1948–49 it took seventeen months to reach a similar position. In the more severe contractions none of the aggregates got back to the peak level in that length of time. Among leading indicators (section B), five of the ten were above the July 1953 peak levels within ten months, or three months before the business upturn in August 1954. In the 1948–49 contraction five of the ten leaders had exceeded the peak level after thirteen months, or two months after the business upturn began. In 1920, 1937, and 1929 none had regained the peak level even after a full year of contraction. In Chart 290 the solid curves depict these developments during the first 24 months of each contraction.

Even though an indicator has not risen sufficiently to exceed the level attained when business activity was at its zenith, it may nevertheless have begun to rise. It is vital to watch for these initial upward movements. The right-hand section of Table 278 (and the dotted curves in Chart 290) shows what propor-

[10] This implies that the leaders indicate the *subsequent* ranking of the aggregates. The following set of rank correlation coefficients, based on the ranks of the average ranks in Table 275, supports that inference, as well as the observation that the ranks of the leaders stabilize at an earlier date:

	Months after Peak		
	3	6	9
Leaders in specified month vs.			
Aggregates 3 mos. later	.88	.95	.99
Aggregates 6 mos. later	.70	.86	—
Leaders 3 mos. later	.85	.88	.76
Aggregates in specified month vs.			
Aggregates 3 mos. later	.29	.59	.83
Aggregates 6 mos. later	−.14	.39	—

TABLE 278

MEASURES OF THE CHANGING SCOPE OF BUSINESS CYCLE CONTRACTIONS

A. TEN INDICATORS OF AGGREGATE ECONOMIC ACTIVITY

Months after Peak	Per Cent Exceeding Level Reached at Business Cycle Peak[a] of								Per Cent Exceeding Level Reached in Third Preceding Month[b]							
	July 1957[c]	Oct. 1926[d]	July 1953	Nov. 1948	May 1923[d]	Jan. 1920[e]	May 1937	Aug. 1929	July 1957[c]	Oct. 1926[d]	July 1953	Nov. 1948	May 1923[d]	Jan. 1920[e]	May 1937	Aug. 1929
1	65	44	20	20	50	42	40	20	80	56	50	40	75	100[f]	60	55
2	20	31	10	0	25	83	75	20	30	31	30	0	44	100[f]	55	20
3	20	44	10	0	25	50	50	10	20	31	0	0	25	25	50	5
4	10	31	0	0	12	50	40	0	0	38	10	0	12	58	25	0
5	10	56	10	10	25	67	10	0	20	56	5	20	38	42	0	10
6	10	25	10	0	12	50	0	0	30	31	5	20	38	75	0	0
7	10	25	0	0	25	67	0	0	0	19	5	10	56	50	0	0
8	10	25	0	0	38	50	0	0	0	12	10	0	81	25	0	10
9		44	0	0	50	25	0	0		25	40	10	100	17	0	10
10		50	0	10	25	17	0	0		50	45	50	75	8	0	0
11		38	10	0*	12	8	0	0		56	65	35*	25	0	0	0
12		31	0	0	12	17	0	0		56	65	40	0	17	0	0
13		44*	10*	0	12	17	0*	0		38*	50*	70	6	0	20*	5
14		44	15	20	0*	8	0	0		44	50	100	6*	8	65	5
15		44	20	30	19	8	0	0		44	60	80	38	0	90	0
16		44	40	40	12	0	0	0		69	100	100	69	25	100	0
17		38	40	50	25	0	0	0		75	100	100	88	25	90	0
18		38	50	60	50	0*	0	0		69	100	100	100	25*	90	5
19		44	50	60	50	0	0	0		88	100	90	94	42	90	35
20		56	60	70	44	0	0	0		69	80	100	100	50	80	40
21		50	60	80	50	0	0	0		62	100	100	81	83	45	10
22		56	70	80	50	0	0	0		62	90	100	88	75	55	5
23		75	80	80	50	0	0	0		88	95	90	56	83	60	0
24		69	80	80	50	0	0	0		88	100	75	38	42	70	0

* Business cycle trough was reached in this month.

[a] Three-month average, centered on the peak month.

[b] The third month after the peak is compared with the peak month, the fourth month after the peak is compared with the first month after the peak, etc. The results are entered in the terminal month of each comparison.

[c] Data for corporate profits for IV 1957 and I 1958 and for GNP for I 1958 are not yet available (March 1958). It is assumed that both series have declined since the peak quarter (III 1957).

[d] Based on 8 out of 10 series; nonagricultural employment and unemployment rate not available.

[e] Based on 6 out of 10 series; nonagricultural employment, unemployment rate, GNP, and personal income not available.

[f] Based on 5 series; corporate profits after taxes start in the first quarter of 1920.

B. TEN LEADING INDICATORS

Months after Peak	Per Cent Exceeding Level Reached at Business Cycle Peak[a] of								Per Cent Exceeding Level Reached in Third Preceding Month[b]							
	July 1957	Oct. 1926	July 1953	Nov. 1948	May 1923	Jan. 1920[g]	May 1937	Aug. 1929	July 1957	Oct. 1926	July 1953	Nov. 1948	May 1923	Jan. 1920[g]	May 1937	Aug. 1929
1	40	30	30	20	30	44	10	20	60	30	10	10	25	38[h]	20	50
2	20	20	10	10	15	22	20	20	30	20	30	20	10	25[h]	20	20
3	20	10	40	0	10	33	20	10	10	30	20	0	30	44	30	0
4	0	20	30	0	20	11	10	10	0	30	40	30	25	11	10	10
5	10	50	40	0	20	11	0	0	10	80	30	25	40	22	0	0
6	0	30	40	20	10	11	0	0	30	50	35	40	50	11	0	40
7	0	20	30	10	30	11	0	0	45	70	55	45	80	11	0	20
8	0	20	40	20	40	0	0	0	38	25	60	50	70	11	20	40
9		10	40	20	50	0	0	0		35	70	80	70	11	30	30
10		20	50	40	20	0	0	0		40	60	90	20	0	50	20
11		30	50	40*	10	0	0	0		40	90	95*	0	0	40	20
12		30	50	40	20	0	0	0		75	80	70	0	11	50	10
13		50*	40*	50	20	0	0*	0		80*	80*	65	10	22	40*	30
14		50	40	50	30*	0	20	0		70	80	90	30*	56	90	20
15		50	60	60	20	0	20	0		90	75	80	50	67	90	30
16		60	40	80	20	0	10	0		60	70	80	70	78	90	10
17		40	60	80	30	0	20	0		40	90	80	90	44	80	20
18		60	60	90	40	0*	20	0		60	90	90	90	33*	90	60
19		60	70	80	50	11	20	0		60	95	100	100	78	75	75
20		60	65	90	50	11	20	0		90	70	80	80	89	55	70
21		70	65	100	50	22	20	0		55	60	80	70	78	50	35
22		60	90	90	50	22	10	0		50	55	80	50	78	75	20
23		40	70	100	50	11	10	0		50	65	70	20	44	10	10
24		65	75	100	50	11	10	0		50	75	20	40	67	30	10

[a] Based on 9 out of 10 series; average workweek not available.

[b] Based on 8 series; new orders of durable goods start January 1920.

The seven business contractions before 1957 are arrayed from left to right in order of their severity. Series that are unchanged from peak level, or from third preceding month, are counted one-half in determining the per cent exceeding peak level or third preceding month. The series included are listed in Tables 266 and 275; in addition, the unemployment rate is included with the ten indicators of aggregate economic activity. A rise in the unemployment rate, layoff rate, or business failures is counted as a decline, and vice versa.

tion of the indicators began to register recovery in that sense as the contractions proceeded. It compares the current monthly figure with the figure three months earlier.[11] We find that in the 1953–54 and 1948–49 episodes substantial proportions of the ten aggregates were rising by the tenth month after the business recessions began (i.e., several months before the end of the contractions). But in 1937–38 and 1929–30 none of the aggregates were rising at this stage, even though in 1937–38 the end of the contraction came in the thirteenth month (June 1938). The recoveries in the leading indicators took place earlier and were more widespread. Even in 1937–38, half of the leaders had registered increases from the three-month-ago level by the tenth month, March 1938, thus heralding the end of that recession.

A sustained narrowing of the scope of the contraction is a signal that it is about to end. For example, the 1937–38, 1948–49, and 1953–54 contractions ended about a year after they began. The evidences of sustained recovery became apparent at an earlier date in the milder 1948–49 and 1953–54 episodes than in the sharp 1937–38 contraction. In each case, however, the narrowing of the scope of the contraction occurred earlier in the leaders than in the aggregates. At times, as in 1923–24, a narrowing of the scope of contraction may be followed by a further widening before recovery finally gets under way.[12]

The measures provided in Table 278, especially those showing the proportion of indicators that have exceeded their levels at the peak, suggest that if a period of depressed activity is defined as the interval from the peak to the date when activity regains the peak level, this period may be more closely related to the severity or magnitude of the decline from peak to trough than the duration of the decline itself is. Of course, different indicators of aggregate economic activity regain their previous peak levels at different dates. If we take as one criterion the date when half of our ten indicators of aggregate activity regained their level of the preceding business cycle peak, and as an alternative estimate the date when the industrial production index did so, we obtain the results shown in Table 281.

Not only are the periods of depressed activity substantially longer, as would be expected, than the contraction durations; they are also more closely related to the severity of the contractions. It took approximately a year and a half for the economy to recover, in the sense defined above, from the four mild or moderate contractions of 1926, 1953, 1948, and 1923; two and a half to three years to recover from the sharp contractions of 1920 and 1937; and at least seven years to recover from the 1929 contraction.

Another way to look at these results is as follows. Except in the Great Depression about a year or year and a half elapsed before the economy generated an upturn in business activity, whether the contraction was severe or moderate. But after the upturn the recovery to the previous peak was attained

[11] The choice of this interval involved a reconciliation of two conflicting considerations. A shorter interval would reflect a cyclical upturn more promptly, but would also reflect erratic movements more frequently. A longer interval would have the opposite advantage and disadvantage. The interval could well be different for different series. (A criterion for selecting an appropriate interval is developed by Julius Shiskin in *Electronic Computers and Business Indicators*, National Bureau of Economic Research, Occasional Paper 57, 1957.)

[12] For an analysis of the timing of changes in the scope of business cycle expansions and contractions and the relation of scope to severity, see the references cited in note 5 and also Geoffrey H. Moore, "The Diffusion of Business Cycles," in *Economics and the Public Interest*, Robert A. Solo, editor (Rutgers University Press, 1955), pp. 35–64.

in a much shorter time after the moderate contractions than after the severe ones. These recoveries were accomplished in 7, 5, 6, and 4 months after the upturn in the 1926, 1953, 1948 and 1923 contractions, respectively (col. 6 minus col. 3, Table 281). Recovery took roughly three times as long, 17 and 18 months, after the upturns in the 1920 and 1937 contractions, and nearly four years in the Great Depression. Although recoveries from severe contractions have usually taken place at a more rapid rate than from moderate contractions (that is, the percentage rate of growth per month is higher), the difference has

TABLE 281

DURATIONS OF BUSINESS CONTRACTIONS AND PERIODS OF DEPRESSED ACTIVITY

Business Cycle Peak Date	Business Cycle Trough Date	Duration of Contraction, Months from (1) to (2)	Date when Peak Level Was Regained[a]		Period of Depressed Activity		Percentage Decline in Industrial Production from Business Cycle Peak to Trough
			Estimate A	Estimate B	Estimate A [months from (1) to (4)]	Estimate B [months from (1) to (5)]	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Oct. 1926	Nov. 1927	13	June 1928[b]	June 1928[c]	20	20	− 5.7
July 1953	Aug. 1954	13	Jan. 1955	May 1955	18	22	− 9.5
Nov. 1948	Oct. 1949	11	Apr. 1950	Apr. 1950	17	17	− 7.7
May 1923	July 1924	14	Nov. 1924[d]	July 1925	18	26	−16.3
Jan. 1920	July 1921	18	Dec. 1922	Oct. 1922	35	33	−29.0
May 1937	June 1938	13	Dec. 1939	Oct. 1939	31	29	−31.5
Aug. 1929	Mar. 1933	43	[e]	Nov. 1936	[e]	87	−50.1

[a] Estimate A is the date when at least half of the following indicators of aggregate economic activity had regained their respective levels at the business cycle peak date: nonagricultural employment, unemployment rate, gross national product, industrial production, freight carloadings, bank debits, personal income, retail sales, corporate profits, wholesale price index. Not all indicators were available for every date. Estimate B is the date when the industrial production index regained the level it had reached at the business cycle peak (three-month average centered on the peak).

[b] Half the indicators were above their peak levels temporarily in March and August 1927, or five and ten months after the peak, respectively, but before the trough was reached.

[c] The production index attained the peak level temporarily in March 1927, five months after the peak but before the trough was reached.

[d] Half the indicators were above their peak levels temporarily in February 1924, nine months after the peak but before the trough was reached.

[e] Industrial production was the only one of the ten indicators that exceeded its level as of the August 1929 peak before the next business cycle peak of May 1937.

not been sufficient to compensate fully for the greater depth of the severe contractions. Hence the recoveries take longer.[13]

All this suggests that our measures of the percentage declines in various indicators in the early stages of contraction may indicate, though only in a very rough way, the duration of "depression" as well as the severity of contraction. No great precision can be expected, however. The course of a contraction and the subsequent recovery is not foreordained. Both can be and have been influenced by deliberate actions taken to shorten the contraction and hasten

[13] This implies a greater degree of uniformity in rates of cyclical expansion than in rates of contraction. For some observations on this point see my introduction to Daniel Creamer's *Personal Income during Business Cycles* (National Bureau of Economic Research, 1956), pp. xxiii–xxviii, and the National Bureau's *Thirty-seventh Annual Report*, May 1957, p. 52.

recovery, as well as by events and policies that originate elsewhere. Historical patterns and relationships ought not to be transplanted mechanically. They can and should be used to help us formulate realistic appraisals of existing situations.

5. OTHER APPLICATIONS

So far, attention has been on the use of the measures provided in this report to indicate the severity of recessions and to judge the prospects of recovery. Another use may be mentioned briefly. By applying the measures to a wide variety of economic data one can determine some of the distinctive characteristics of each recession—what sectors or aspects of economic activity are strong and what are weak. For example, in both the 1953–54 and 1948–49 recessions residential building displayed great strength; in 1937–38 and 1929–30, notable weakness. These differences are sharply etched in the percentage changes in the volume of contracts (Appendix A). Six months after recession began, residential contract volume (seasonally adjusted) had dropped 40 per cent in 1929 and 21 per cent in 1937, but had risen 3 per cent in 1948–49 and 10 per cent in 1953–54. Consumer instalment credit advanced vigorously in 1948–49 (it was 28 per cent higher a year after the recession began), rose moderately in 1953–54 (4 per cent higher), but declined appreciably in 1929–30 (9 per cent lower), and 1937–38 (6 per cent lower). Such differentiation of the strong and weak sectors in the economy during a recession is essential to appropriate diagnosis and prescription of policies to encourage revival.

For this purpose, too, the type of measure presented here may well be extended to other data of strategic interest from a policy standpoint. Data on comparative changes in personal and corporate income tax payments, in unemployment benefits, in federal and in state and local expenditures, in public works contracts, in interest rates, in the money supply, and in Federal Reserve operations would enable one to appraise the strength and timing of either deliberate or "built-in" stabilization policies. The simple technique illustrated in this report can thus be adapted to provide an up-to-date, objective set of facts on which to judge not only the severity, scope, and unique character of a developing recession, but also the prospects for an early recovery and the vigor with which steps are being taken to bring recovery about.

In order to facilitate the application of this analysis an electronic computer program has been prepared for the IBM 704 which computes the percentage changes from peak month to the first, second, third and up to the twenty-fourth month after peak, for any given series and for any given list of peak dates. The program also makes a similar set of computations of percentage changes from trough months, so that it can be used to compare cyclical revivals as well as recessions. In addition, the total percentage change from peak to trough and trough to peak is computed, to provide a measure of the full amplitude of cyclical swings. Appendix A is, in fact, a photocopy of the print-out provided by the program (total peak to trough changes are omitted for lack of space).

6. 1957–1958 RECESSION

How does the current contraction in business activity compare with earlier contractions when measured by the methods described above? As already noted,

such a comparison requires that the date of the peak from which the contraction began be established, at least on a tentative basis. Accordingly, July 1957 was selected as the monthly business cycle peak date, and the third quarter of 1957 as the quarterly peak date. This determination was made in October 1957, when data through September were available for the more important economic series.

The selection of a peak date was difficult because a moderate decline in output and employment in manufacturing had been under way since late in 1956, while activity in most other sectors continued to advance. Indeed, some important factors, such as total personal income, had scarcely begun to decline by September 1957. Because of the continued rise in prices and wages, physical volume series on output, income, and trade reached peaks earlier than the corresponding value series. In general, during the first half of 1957 the physical volume of labor input (manhours) and output of the economy at large remained nearly constant or at best gently rising, while greater increases occurred in the pecuniary volume of output, trade, and income. Although declines in activity became widespread between August and September and have been extended since, it is difficult to say whether July or August should be considered the zenith. Subsequent revisions of the data may shift the weight of evidence to August, or, less likely, to an earlier month. It must be observed, too, that at the time the peak date was selected, the contraction had not yet become sufficiently pronounced, or lasted long enough, to qualify as a business cycle contraction in the National Bureau's chronology. The judgment that it was likely to become so has, of course, since been validated by events.

Once the peak date was determined, we could construct the tables in Appendix A that show the extent of the decline after the July 1957 peak for many economic series, and compare this decline with what occurred during similar intervals of time in previous business contractions. The first comparisons were with the two relatively moderate postwar contractions, 1948–49 and 1953–54, and the two severe prewar contractions that began in 1929 and 1937. Later, with the aid of the computational program developed for the IBM 704 electronic computer, we extended the comparisons to earlier cycles and made the computations in several variant forms.

Appendix A presents the record of what these comparisons have shown about the current recession as it developed, and how it stood at the time this report went to press. Drawing upon that record, Table 284 compares the percentage changes for the first seven months of the current recession, i.e., July 1957 to February 1958, with those for the corresponding periods of the preceding recessions. The summary columns at the right show how many of the indicators experienced a smaller decline in this recession than in the earlier ones, and how many a larger decline. Both the indicators of aggregate activity and the leading indicators are decisive in recording a substantially smaller decline in 1957–58 than in the severe contractions of 1937–38 and 1929–30. In comparison with the severe contraction of 1920–21 the leading indicators also show a substantially smaller decline, while most of the aggregates show a larger decline currently. Historically, the leaders have been a more reliable guide at this stage (seven months after the peak), and hence are entitled to greater weight. Thus, our evidence points to a less severe contraction on this occasion than those experienced in 1920–21, 1929–30, or 1937–38.

TABLE 284

PERCENTAGE CHANGES DURING FIRST *SEVEN* MONTHS OF EIGHT BUSINESS CYCLE CONTRACTIONS, TWO GROUPS OF INDICATORS

A. TEN INDICATORS OF AGGREGATE ECONOMIC ACTIVITY

Business Cycle Contraction, First Seven Months	Nonagricultural Employment	Unemployment Rate[a]	Gross National Product (1st 2 quarters)	Industrial Production	Freight Carloadings	Bank Debits Outside N.Y.C.	Personal Income	Retail Sales	Corporate Profits after Taxes (1st 2 quarters)	Wholesale Price Index	Number Showing that 1957–1958 Decline is	
											Smaller	Larger
	Percentage Change											
July 1957–Feb. 1958	−3.4	(+2.4)	−3.4	−10.3	−17.8	−4.7	−1.3	−4.9	n.a.	+0.1	–	–
Oct. 1926–May 1927	n.a.	n.a.	+1.1	−1.3	−4.1	+3.9	n.a.	−2.7	−14.7	−6.0	1	5
July 1953–Feb. 1954	−2.3	(+2.7)	−2.7	−8.3	−11.6	−0.8	−1.0	−2.5	−26.4	−0.1	2	7
Nov. 1948–June 1949	−3.3	(+1.9)	−2.6	−8.4	−15.4	−5.9	−3.1	−0.7	−16.9	−5.0	3	6
May 1923–Dec. 1923	n.a.	n.a.	−1.0	−8.2	−2.9	−4.1	n.a.	−1.0	−39.9	−7.4	1	5
Jan. 1920–Aug. 1920	n.a.	n.a.	n.a.	−3.9	+1.6	+0.7	n.a.	+6.5	−34.1	+14.2	0	5
May 1937–Dec. 1937	−4.4	n.a.	−7.8	−27.3	−24.6	−8.1	−8.1	−9.2	−52.2	−3.0	8	0
Aug. 1929–Mar. 1930	−5.4	n.a.	−5.5	−12.6	−11.1	−16.5	−8.1	−3.5	−41.9	−3.6	6	2
	Ranking of Percentage Change										Average Rank, Five Indicators[c]	Rank of Average Rank
July 1957–Feb. 1958	2	2	5	6	7	5	2	7	—	2	4.0	5.5
Oct. 1926–May 1927	—	—	1	1	3	1	—	5	—	7	1.0	1
July 1953–Feb. 1954	1	3	4	4	5	3	1	4	—	3	2.6	3
Nov. 1948–June 1949	3	1	3	5	6	6	3	2	—	6	4.0	5.5
May 1923–Dec. 1923	—	—	2	3	2	4	—	3	—	8	3.0	4
Jan. 1920–Aug. 1920	—	—	—	2	1	2	—	1	—	1	2.0	2
May 1937–Dec. 1937	4	—	7	8	8	7	4	8	—	4	6.0	7
Aug. 1929–Mar. 1930	5	—	6	7	4	8	5	6	—	5	6.2	8
Rank Correlation Coefficient[b]	1.00	—	.71	.64	.07	.75	1.00	.29	.89	−.29	—	.75

[a] Change in rate, not percentage change in rate. Unemployment rate and layoff rate are ranked invertedly, with smallest increase given a rank of 1.

[b] See Tables 273 and 274. Coefficients are based on contractions prior to 1957–58.

[c] Nonagricultural employment, gross national product, industrial production, bank debits, personal income; corporate profits omitted because data for first quarter of 1958 not available.

B. TEN LEADING INDICATORS

Business Cycle Contraction, First Seven Months	Average Workweek	Layoff Rate[a]	Gross Accession Rate[a]	New Orders, Durable Goods	Residential Construction Contracts	Comm. & Indus. Construction Contracts	New Incorporations	Business Failures, Liabilities	Basic Commodity Prices	Industrial Stock Prices	Number Showing that 1957–1958 Decline is	
											Smaller	Larger
	Percentage Change											
July 1957–Feb. 1958	−3.9	(+1.2)	(−0.6)	−19.7	−23.3	−21.2	−6.5	+14.7	−4.6	−11.6	—	—
Oct. 1926–May 1927	−0.5	(+0.1)	(−0.5)	+5.7	−20.3	−1.6	−0.5	+13.1	−2.5	+8.2	0	10
July 1953–Feb. 1954	−2.5	(+1.1)	(−1.2)	−11.5	+9.7	−20.8	+4.5	+16.1	−0.04	+8.3	2	8
Nov. 1948–June 1949	−2.2	(+0.8)	(−0.4)	−16.6	+31.2	−14.8	−1.6	+13.5	−25.4	−8.6	1	9
May 1923–Dec. 1923	−3.2	(+0.1)	(−6.3)	+27.5	+11.9	−5.2	−4.7	−5.3	−1.2	−2.2	1	9
Jan. 1920–Aug. 1920	n.a.	(+0.3)	(+0.5)	−22.2	−46.6	−63.6	−16.1	+283.3	−10.6	−17.0	7	2
May 1937–Dec. 1937	−14.6	(+4.1)	(−1.3)	−37.7	−33.4	−36.2	−11.2	+99.2	−24.9	−28.2	10	0
Aug. 1929–Mar. 1930	−7.1	(+0.5)	(−2.3)	−29.8	−44.0	−27.2	−10.0	+86.0	−11.3	−21.5	9	1
	Ranking of Percentage Change										Average Rank, Ten Indicators	Rank of Average Rank
July 1957–Feb. 1958	5	7	4	5	5	5	5	4	4	5	4.9	5
Oct. 1926–May 1927	1	1	3	2	4	1	2	2	3	2	2.1	1
July 1953–Feb. 1954	3	6	5	3	3	4	1	5	1	1	3.2	3
Nov. 1948–June 1949	2	5	2	4	1	3	3	3	8	4	3.5	4
May 1923–Dec. 1923	4	2	8	1	2	2	4	1	2	3	2.9	2
Jan. 1920–Aug. 1920	n.a.	3	1	6	8	8	8	8	5	6	5.9	6
May 1937–Dec. 1937	7	8	6	8	6	7	7	7	7	8	7.1	8
Aug. 1929–Mar. 1930	6	4	7	7	7	6	6	6	6	7	6.2	7
Rank Correlation Coefficient[b]	.89	.32	.36	.75	.57	.71	.82	.61	.46	.89	—	.86

On the other hand, nearly all the leading series, and a majority of the aggregate indicators, registered larger declines in the seven months since July 1957 than in the corresponding periods of each of the four milder business contractions, namely, 1926–27, 1953–54, 1948–49, and 1923–24. The 1957–58 contraction appears, therefore, to be of intermediate proportions, not the mildest on record nor yet the most severe (see Chart 287). The rankings shown in the bottom section of the table confirm this indication.[14]

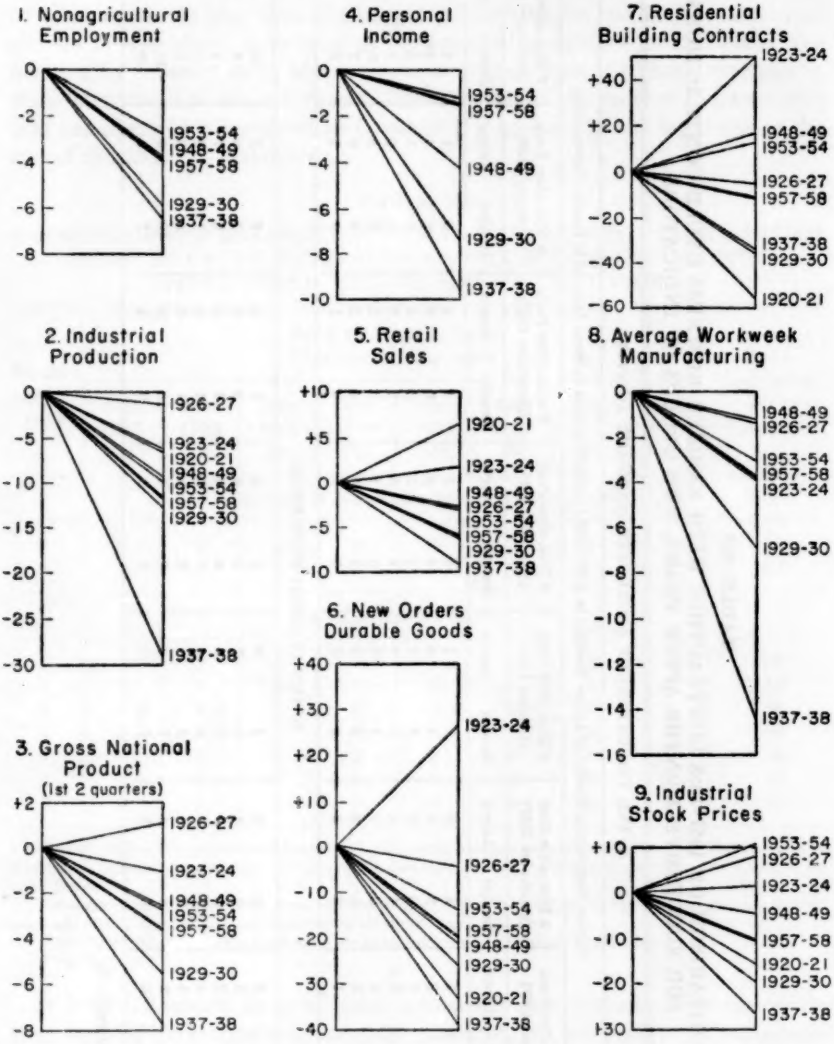
This evaluation of the severity of the 1957–58 contraction began emerging from the leading series when data became available for the fourth month after the July peak, i.e. for November 1957 (Tables 288 and 289). As we have seen in the preceding pages, a historical ranking of the recessions based on changes in the leading series during the first four months accords fairly well with their ultimate ranking according to severity. Hence when it became apparent (in late December) from data covering the four-month span July–November 1957 that declines in most of the leading series were larger than in the milder recessions in our list but smaller than in the most severe recessions, this bit of evidence helped to support other indications pointing to a recession of this general character. Data for five-, six-, seven and eight-month spans have provided further support.

On the other hand, the historical analysis of the indicators of aggregate economic activity suggested that a ranking of the recessions even moderately consistent with their ultimate severity would not emerge until at least six months had elapsed. Thus a comparison of the declines in these aggregates in the current recession with their declines in earlier recessions was not likely to yield consistent results during the first six months. This was indeed the case. Changes in the aggregates over a six-month span showed that the current decline was substantially smaller than in the severe contractions of 1929 and 1937, and slightly smaller than in the mild contractions of 1926 and 1953. But most of them also showed *larger* declines than in the contractions of 1923 and 1920, which rank fourth and fifth respectively in our list according to severity. This inconsistency is not likely to be erased until data for the aggregative indicators for the first nine months (i.e., through April 1958) are available. If the leading series turn out to be a reliable guide in this respect, the aggregates should then show larger declines than in the 1926–27, 1953–54, 1948–49, and 1923–24 contractions but smaller declines than in the 1920–21, 1937–38, and 1929–30 contractions.

The business contraction that began in mid-1957 rapidly engulfed most of our indicators, as its predecessors had done also. Table 278 shows that by October, the third month after the peak, eight of the ten aggregates and eight of the ten leading series had moved down from their levels at the peak. The

[14] If August 1957 were selected as the business cycle peak instead of July, the current contraction would appear somewhat more severe relative to the earlier contractions, because the six-month decline from August to February in many of the indicators of aggregate activity would be nearly the same as the seven-month decline from July to February (since the three-month averages centered on July and August would be nearly alike), but the decline would be compared with six-month declines in the earlier contractions instead of with seven-month declines. The average ranks for the aggregate indicators in Table 284 would be only slightly altered, however, and those for the leading series scarcely changed at all. Thus the shift would not affect the conclusions given above.

CHART 287. Comparative changes during first eight months of business cycle contractions, selected indicators.



Percentage changes are computed from the three-month average of each indicator centered on the business cycle peak. The eight-month intervals are: Jan. 1920–Sept. 1920, May 1923–Jan. 1924, Oct. 1926–June 1927, Aug. 1929–Apr. 1930, May 1937–Jan. 1938, Nov. 1948–July 1949, July 1953–Mar. 1954, July 1957–Mar. 1958.

wide scope of the decline was maintained in succeeding months. More recently, a slight narrowing of the scope of the contraction has appeared, especially in the leading series, when the measurement is made in terms of changes over a

TABLE 288

SUMMARY COMPARISON OF 1957–1958 CONTRACTION WITH EARLIER BUSINESS CYCLE CONTRACTIONS, FOR SUCCESSIVE MONTHS AFTER PEAKS, TWO GROUPS OF INDICATORS

A. TEN INDICATORS OF AGGREGATE ECONOMIC ACTIVITY

Earlier Contraction, Starting in	Number Registering Smaller or Larger Declines in 1957–1958 Contraction than in Same Interval in Earlier Contractions													
	1 Mo. after Peak (August 1957)		2 Mos. after Peak (September 1957)		3 Mos. after Peak (October 1957)		4 Mos. after Peak (November 1957)		5 Mos. after Peak (December 1957)		6 Mos. after Peak (January 1958)		7 Mos. after Peak (February 1958)	
	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger	Smaller	Larger
Oct. 1926	4	1	2	3	2	5	1	4	1	4	3	4	1	4
July 1953	6	2	5	3	4	5	4	4	5	3	5	4	2	6
Nov. 1948	7	1	6	2	7	2	6	2	6	2	6	3	3	5
May 1923	2.5	2.5	4	1	2	5	2	3	2	3	1	6	1	4
Jan. 1920	3	2	0	5	2	3	0	5	0	5	0	5	0	5
May 1937	5	2	1	6	1	7	2	5	5	2	8	0	7	0
Aug. 1929	6	1	5	2	6	2	5	2	6	1	7	1	5	2
B. TEN LEADING INDICATORS														
Oct. 1926	5	5	5	5	4	6	2	8	1	9	3	7	0	10
July 1953	5	5	6	4	3	7	2	8	5	5	1	9	2	8
Nov. 1948	8	2	8	2	5	5	4	6	4	6	3	7	1	9
May 1923	6	4	7	3	4	6	5	5	4	6	1	9	1	9
Jan. 1920	5	4	5	4	5	4	6	3	4	5	6	3	7	2
May 1937	8	2	6	4	7	3	7	3	10	0	10	0	10	0
Aug. 1929	7	3	5	5	7	3	7	3	9	1	7	3	9	1

Source: Table 284 and Appendix A.

three-month span (see Chart 290).[15] The improvement has not yet carried as far as the similar improvement had at the end of the first six months of 1953–54 contraction. At that time (January 1954) four of the ten leading series had already moved above their level at the peak, whereas none had achieved this position by January 1958. Moreover, as is evident from the chart, reversals in these measures are not infrequent. Nevertheless, if the modest improvement that has occurred is sustained and extended, it may signal the beginning of the end of the 1957–58 contraction.

TABLE 289

RANKING OF 1957–1958 CONTRACTION IN RELATION TO SEVEN EARLIER BUSINESS CYCLE CONTRACTIONS, FOR SUCCESSIVE MONTHS AFTER PEAKS, TWO GROUPS OF INDICATORS

Months after Peak	Rank of Average Rank Contraction beginning								Rank Correlation with Severity[b]
	July 1957 (1)	Oct. 1926 (2)	July 1953 (3)	Nov. 1948 (4)	May 1923 (5)	Jan. 1920 (6)	May 1937 (7)	Aug. 1929 (8)	
Six Indicators of Aggregate Economic Activity[a]									
3	6	2	4	8	3	5	1	7	.18
6	4.5	1	4.5	6	3	2	7	8	.68
Ten Leading Indicators[c]									
1	1	2	3	7	4.5	6	8	4.5	.32
2	1	2	3	8	5	7	6	4	.50
3	3.5	1.5	1.5	6	3.5	5	7	8	.88
4	5	1	2	3	4	7	6	8	.96
5	4	1	2	3	5	6	8	7	.96
6	5	1	3	4	2	6	8	7	.86
7	5	1	3	4	2	6	8	7	.86

Source: Tables 275 and 284, and Appendix A. The seven business contractions before 1957–1958 are arrayed from left to right according to their over-all severity (see text). A rank of 1 indicates a small decline, a rank of 8 a large decline.

[a] Nonagricultural employment, gross national product, industrial production, bank debits, personal income, corporate profits. Not all series are available for all cycles—see Table 284.

[b] See Table 275. Correlation coefficients are based on contractions prior to 1957–58.

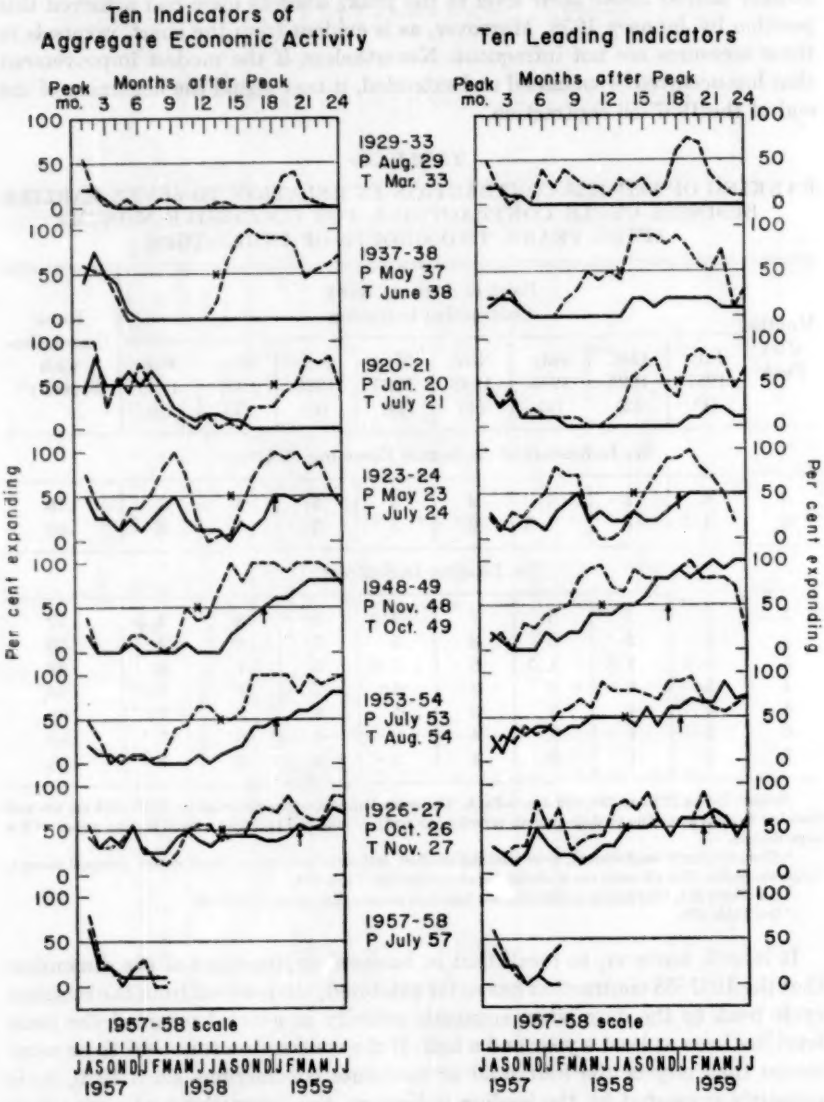
[c] See Table 276.

It is well, however, to recall that in business contractions of the dimension that the 1957–58 contraction has so far exhibited, the interval from the business cycle peak to the date when economic activity in general regained the peak level has been at least a year and a half. If the contraction turns out to be more severe than any of the four mild or moderate contractions since 1920, as is presently suggested by the leading indicators, this interval may be exceeded.

[15] Julius Shiskin pointed out (in February 1958) a similar development in several diffusion indexes compiled at the Bureau of the Census for the Council of Economic Advisers, and it has appeared in some other diffusion indexes compiled at the National Bureau of Economic Research. These indexes show, for example, that there has been a slight increase in the number of manufacturing industries reporting an advance in the workweek since autumn 1957, and a slight increase in the number reporting a rise in new orders.

CHART 290. Measures of the scope of business cycle contractions.

——— Per cent exceeding level reached at business cycle peak
– – – – – Per cent exceeding level reached in third preceding month



Asterisks represent business cycle troughs.

Arrows represent dates when half of ten indicators of aggregate economic activity exceeded level reached at business cycle peak. For 1920–21 this date is December 1922, 35 months after the January 1920 peak; for 1937–38 it is December 1939, 31 months after the May 1937 peak (see Table 278).

On the other hand, these measures do not suggest that the period of less than peak activity will extend to the two and a half years that characterized the severe contractions of 1920–21 and 1937–38. In any event, a great deal depends on the governmental and private actions undertaken to bring about an early and rapid recovery.

Some of the distinctive characteristics of the 1957–58 contraction, as it has developed so far, may be observed in Table 284. One of the outstanding facts is the relatively small decline in personal income. The decline of 1.3 per cent, July 1957 to February 1958, is only one-sixth as large as the 8.1 per cent declines during the first seven months of the 1937 and 1929 contractions. Yet the current decline in employment is more than half as large as in 1937 and 1929, and the current decline in gross national product is also a substantial fraction of the corresponding declines in 1937 and 1929. The greater stability of personal income is partly attributable to the larger role of unemployment compensation and other transfer payments, partly to the growth of employment in government, distributive trades, and other service industries where rates of pay and the volume of employment are typically more stable.[16] In terms of disposable personal income (that is, income after federal income taxes), the contrast between the current recession and those of the thirties would be even greater.

Another unusual and perhaps not unrelated feature of the 1957–58 recession is the fact that the wholesale price index after eight months of recession is still above its level when the recession began. Except for 1920, this is the only recession in which this index, which covers all commodities except farm products and foods, has not declined, although in 1953–54 the decline during the first eight months was minute. In the case of the consumers price index, the rise during the current recession is less exceptional, as the following figures indicate:

Business Cycle Contraction, beginning	Percentage Change during					
	First Eight Months			First Twelve Months		
	Basic Commodity Price Index	Wholesale Price Index, Excl. Farm & Food	Consumers' Price Index	Basic Commodity Price Index	Wholesale Price Index, Excl. Farm & Food	Consumers' Price Index
July 1957	−4.6	+0.1	+2.6			
Oct. 1926	−2.8	−6.0	+0.6	+4.5	−5.9	−1.4
July 1953	+2.0	−0.3	+0.4	+4.2	−0.2	+0.3
Nov. 1948	−24.3	−5.3	−2.1	−22.1	−5.1	−1.9
May 1923	−2.1	−4.7	+1.2	−8.1	−7.0	+0.4
Jan. 1920	−15.0	+9.7	+3.8	−39.2	−18.6	−1.2
May 1937	−23.9	−3.2	−0.9	−32.4	−5.3	−1.7
Aug. 1929	−12.4	−4.0	−0.7	−17.4	−8.6	−4.4

[16] Cf. Daniel Creamer, *Personal Income during Business Cycles* (National Bureau of Economic Research, 1956). Another factor is the secular decline in the proportion of total income derived from farming, which has in the past declined more sharply than total nonfarm income during business recessions. Indeed, in the current recession farm income has been increasing rather than declining. The increasing number of wage contracts that tie rates of pay to the cost of living, which has risen in the first eight months of this recession, may have imparted some stability to incomes currently, although this hinges on the question whether the resulting increases in wage rates may not have brought about some reduction in hours or in employment.

On the other hand, the price index of basic commodities (cotton, wool, copper, steel scrap, etc.) has declined during the first eight months of the current recession as it has in every other recession except 1953–54.

In the more severe recessions, all three price indexes declined further in the first twelve months than they had in the first eight. But in the milder recessions an interesting gradation appears. The basic commodity price index showed a smaller decline or an actual increase after twelve months in the 1926–27, 1953–54, and 1948–49 recessions. The decline in the wholesale price index remained about the same. The consumers' price index showed further weakness, so that the declines after twelve months were somewhat greater than after eight (or the increases were less), even in the mild recessions. Thus our figures reflect some of the lagging relations among prices, as well as the influence of mild or deep recessions upon the entire price structure.

Other facets of the 1957–58 contraction are well worth study and reflection when set against the corresponding pattern of events in previous contractions. The "forward look" that these previous contractions provide is illuminating. We have been able to touch on only a few of the many strategic economic variables and relationships that can usefully be analyzed during the course of a business contraction, a fact that underlines the tentative and preliminary character of the experiments reported above.

APPENDIX A

ACTUAL IBM 704 PRINT-OUT SHOWING STANDINGS AND PERCENTAGE CHANGES AFTER BUSINESS CYCLE PEAKS, TWENTY ECONOMIC INDICATORS

NOTE: All series except Wholesale Prices ex Farm Products and Food, Basic Commodity Prices, and Index of Industrial Stock Prices are adjusted for seasonal variations. The peak standings are three-month averages centered on the business cycle peak month. Quarterly data are shown as of the mid-month of the quarter, and peak standings are for the business cycle peak quarter. The units in which the standings are expressed are given at the end of the table, together with brief source notes. An asterisk at the peak standing indicates the beginning of a new segment of the series, identified in the notes. For sources of current data and current seasonal factors see Appendix B.

NONAGRICULTURAL EMPLOYMENT

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1929 8	3090.3	3082	3067	3044	3004	2969	2948	2923	2910	2890	2861	2831	2792
1937 5	2967.7	2974	2988	2987	2974	2948	2901	2838	2778	2757	2733	2718	2679
1945 2	41429.7*	41382	41105	40839	40731	40447	40047	38056	38034	38425	38582	39320	38836
1948 11	44678.3	44559	44206	43949	43726	43554	43342	43216	43030	43107	43321	42455	42792
1953 7	49875.0	49842	49695	49636	49344	49156	48859	48714	48506	48407	48271	48274	48140
1957 7	52810.7	52844	52662	52469	52218	51980	51709	51035	50732				

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1929 8	3090.3	2765	2737	2710	2685	2664	2647	2629	2625	2606	2581	2558	2527
1937 5	2967.7	2665	2672	2692	2713	2729	2758	2784	2789	2795	2799	2766	2793
1945 2	41429.7*	39896	40428	40862	41298	41705	42195	42437	42588	42855	42899	43138	43228
1948 11	44678.3	43083	43003	42703	43333	43707	44068	44558	44992	45589	45899	46207	46291
1953 7	49875.0	48149	48197	48348	48600	48756	48827	48909	49232	49461	49824	50181	50295
1957 7	52810.7												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1929 8	3090.3	-.27	-.75	-1.50	-2.79	-3.93	-4.61	-5.41	-5.84	-6.48	-7.42	-8.39	-9.65
1937 5	2967.7	.21	.69	.65	.21	-.66	-2.25	-4.37	-6.39	-7.10	-7.91	-8.41	-9.73
1945 2	41429.7*	-.11	-.78	-1.43	-1.69	-2.37	-3.34	-8.14	-8.20	-7.25	-6.87	-5.09	-6.26
1948 11	44678.3	-.27	-1.06	-1.63	-2.13	-2.52	-2.99	-3.27	-3.69	-3.52	-3.04	-4.98	-4.22
1953 7	49875.0	-.07	-.36	-.48	-1.06	-1.44	-2.04	-2.33	-2.74	-2.94	-3.22	-3.21	-3.48
1957 7	52810.7	.06	-.28	-.65	-1.12	-1.57	-2.09	-3.36	-3.94				

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1929 8	3090.3	-10.53	-11.43	-12.31	-13.12	-13.80	-14.35	-14.93	-15.06	-15.67	-16.48	-17.23	-18.23
1937 5	2967.7	-10.20	-9.96	-9.29	-8.58	-8.04	-7.06	-6.19	-6.02	-5.82	-5.68	-6.79	-5.89
1945 2	41429.7*	-3.70	-2.42	-1.37	-.32	.66	1.85	2.43	2.80	3.44	3.55	4.12	4.34
1948 11	44678.3	-3.57	-3.75	-4.42	-3.01	-2.17	-1.37	-.27	.70	2.04	2.73	3.42	3.61
1953 7	49875.0	-3.46	-3.36	-3.06	-2.56	-2.24	-2.10	-1.94	-1.29	-.83	-.10	.61	.84
1957 7	52810.7												

UNEMPLOYMENT RATE

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
		STANDINGS					FIRST YEAR AFTER PEAK						
1929 8	58.0	91	231	189	208	221	312	363	360	339	377	502	664
1937 5	1152.7	1150	1097	1114	1163	1219	1370	1591	1688	1788	1849	1896	1974
1948 11	396.0*	418	446	473	492	518	622	589	643	674	650	779	889
1953 7	262.7	263	290	308	354	451	484	530	571	584	597	537	566
1957 7	422.7	429	453	470	493	503	576	666	699				

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
		STANDINGS					SECOND YEAR AFTER PEAK						
1929 8	58.0	779	901	1078	1189	1118	1167	1178	1194	1230	1302	1384	1501
1937 5	1152.7	2000	1958	1918	1860	1822	1714	1636	1761	1743	1719	1739	1690
1948 11	396.0*	680	665	644	627	584	568	528	500	448	445	410	426
1953 7	262.7	601	612	573	535	499	494	482	470	479	428	414	405
1957 7	422.7												

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
		ABSOLUTE CHANGES					FIRST YEAR AFTER PEAK						
1929 8	58.0	33.0	173.0	131.0	150.0	163.0	254.0	305.0	302.0	281.0	319.0	444.0	606.0
1937 5	1152.7	-2.7	-55.7	-38.7	10.3	66.3	217.3	438.3	535.3	635.3	696.3	743.3	821.3
1948 11	396.0*	22.0	50.0	77.0	96.0	122.0	226.0	193.0	247.0	278.0	254.0	383.0	273.0
1953 7	262.7	.3	27.3	45.3	91.3	188.3	221.3	267.3	308.3	321.3	334.3	274.3	303.3
1957 7	422.7	6.3	30.3	47.3	70.3	80.3	153.3	243.3	276.3				

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
		ABSOLUTE CHANGES					SECOND YEAR AFTER PEAK						
1929 8	58.0	721.0	843.0	1020.0	1131.0	1060.0	1109.0	1120.0	1136.0	1172.0	1244.0	1326.0	1443.0
1937 5	1152.7	847.3	805.3	765.3	707.3	669.3	561.3	483.3	608.3	590.3	566.3	586.3	537.3
1948 11	396.0*	284.0	269.0	248.0	231.0	188.0	172.0	132.0	104.0	52.0	49.0	14.0	30.0
1953 7	262.7	338.3	349.3	310.3	272.3	236.3	231.3	219.3	207.3	216.3	165.3	151.3	142.3
1957 7	422.7												

GROSS NATIONAL PRODUCT

DATE	AT PEAK	STANDINGS + 3 MO	FIRST YEAR AFTER PEAK + 6 MO	+ 9 MO	+12 MO
1923 5	8284.0	8304	8204	8512	8448
1926 8	9284.0	9612	9384	9328	9412
1929 8	10556.0	10300	9972	9752	9260
1937 5	9252.0	9720	8928	8164	8148
1945 2	2212.0 *	2237	2122	1971	1967
1948 11	2640.0	2599	2572	2565	2535
1953 5	3674.0	3663	3575	3581	3587
1957 8	4400.0	4326	4260		

DATE	AT PEAK	STANDINGS +15 MO	SECOND YEAR AFTER PEAK +18 MO	+21 MO	+24 MO
1923 5	8284.0	8092	8288	8900	8644
1926 8	9284.0	9316	9908	9248	9712
1929 8	10556.0	8368	8068	7948	7908
1937 5	9252.0	8404	8856		
1945 2	2212.0 *	2050	2152	2201	2232
1948 11	2640.0	2649	2759	2944	3050
1953 5	3674.0	3600	3677	3790	3877
1957 8	4400.0				

DATE	AT PEAK	PER CENT CHANGES + 3 MO	FIRST YEAR AFTER PEAK + 6 MO	+ 9 MO	+12 MO
1929 5	8284.0	.24	-.97	2.75	1.98
1926 8	9284.0	3.53	1.08	.47	1.38
1929 8	10556.0	-2.42	-5.53	-7.62	-12.28
1937 5	9252.0	5.06	-7.82	-11.76	-11.93
1945 2	2212.0 *	1.13	-4.07	-10.89	-11.08
1948 11	2640.0	-1.55	-2.58	-2.84	-3.22
1953 5	3674.0	-.30	-2.69	-2.53	-2.37
1957 8	4400.0	-1.68	-3.41		

DATE	AT PEAK	PER CENT CHANGES +15 MO	SECOND YEAR AFTER PEAK +18 MO	+21 MO	+24 MO
1923 5	8284.0	-2.32	.05	7.44	4.35
1926 8	9284.0	.34	6.72	-.39	4.61
1929 8	10556.0	-20.73	-23.57	-24.71	-28.87
1937 5	9252.0	-9.17	-4.28		
1945 2	2212.0 *	-7.32	-2.71	-.50	.90
1948 11	2640.0	.34	4.51	11.51	15.53
1953 5	3674.0	-2.01	.08	3.16	5.53
1957 8	4400.0				

INDUSTRIAL PRODUCTION

DATE	AT PEAK	STANDINGS + 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	FIRST YEAR AFTER PEAK + 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	42.7	44	43	41	42	42	41	41	40	39	35	33	31
1923 5	49.0	49	48	48	47	46	46	45	46	47	46	45	43
1926 10	52.7	52	52	52	52	53	52	52	52	51	51	50	50
1929 8	60.7	60	59	56	54	54	54	53	53	51	50	48	46
1937 5	64.7	64	65	64	62	57	52	47	46	45	45	44	43
1945 2	124.7	125	122	119	117	111	99	89	86	89	87	85	81
1948 11	103.7	102	101	99	97	96	95	95	94	96	97	94	96
1953 7	136.3	136	133	132	129	126	125	125	123	123	125	124	123
1957 7	145.0	145	144	142	139	135	133	130	128				

DATE	AT PEAK	STANDINGS +13 MO	+14 MO	+15 MO	+16 MO	+17 MO	SECOND YEAR AFTER PEAK +18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	42.7	31	30	30	31	30	30	31	31	33	33	33	34
1923 5	49.0	41	40	42	44	44	45	47	48	48	48	48	48
1926 10	52.7	49	50	51	51	52	51	52	53	53	54	55	56
1929 8	60.7	46	45	44	43	42	42	43	43	43	42	41	40
1937 5	64.7	44	46	49	50	51	54	54	55	54	54	53	53
1945 2	124.7	89	87	84	90	91	94	95	98	97	96	99	99
1948 11	103.7	98	100	99	102	106	110	112	115	120	120	121	120
1953 7	136.3	123	124	126	128	130	132	133	135	136	138	139	139
1957 7	145.0												

DATE	AT PEAK	PER CENT CHANGES + 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	FIRST YEAR AFTER PEAK + 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	42.7	3.13	.78	-3.91	-1.56	-1.56	-3.91	-3.91	-6.25	-8.59	-17.97	-22.66	-27.34
1923 5	49.0	.00	-2.04	-2.04	-4.08	-6.12	-6.12	-8.16	-6.12	-4.08	-6.12	-8.16	-12.24
1926 10	52.7	-1.27	-1.27	-1.27	-1.27	.63	-1.27	-1.27	-1.27	-3.16	-3.16	-5.06	-5.06
1929 8	60.7	-1.10	-2.75	-7.69	-10.99	-10.99	-10.99	-12.64	-12.64	-15.93	-17.58	-20.88	-24.18
1937 5	64.7	-1.03	.52	-1.03	-4.12	-11.86	-19.59	-27.32	-28.87	-30.41	-30.41	-31.96	-33.50
1945 2	124.7	.27	-2.14	-4.54	-6.15	-10.96	-20.69	-28.61	-31.02	-28.61	-30.21	-31.82	-35.03
1948 11	103.7	-1.61	-2.57	-4.50	-6.43	-7.40	-8.36	-8.36	-9.32	-7.40	-6.43	-9.32	-7.40
1953 7	136.3	-.24	-2.44	-3.18	-5.38	-7.58	-8.31	-8.31	-9.78	-9.78	-8.31	-9.05	-9.78
1957 7	145.0	.00	-.69	-2.07	-4.14	-6.90	-8.28	-10.34	-11.72				

DATE	AT PEAK	PER CENT CHANGES +13 MO	+14 MO	+15 MO	+16 MO	+17 MO	SECOND YEAR AFTER PEAK +18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	42.7	-27.34	-29.69	-29.69	-27.34	-29.69	-29.69	-27.34	-27.34	-22.66	-22.66	-22.66	-20.31
1923 5	49.0	-16.33	-18.37	-14.29	-10.20	-10.20	-8.16	-4.08	-2.04	-2.04	-2.04	-2.04	-2.04
1926 10	52.7	-6.96	-5.06	-3.16	-3.16	-1.27	-3.16	-1.27	.63	.63	2.53	4.43	6.33
1929 8	60.7	-24.18	-25.82	-27.47	-29.12	-30.77	-30.77	-29.12	-29.12	-29.12	-30.77	-32.42	-34.07
1937 5	64.7	-31.96	-28.87	-24.23	-22.68	-21.13	-16.49	-16.49	-14.95	-16.49	-16.49	-18.04	-18.04
1945 2	124.7	-28.61	-30.21	-32.62	-27.91	-27.00	-24.60	-23.80	-22.99	-22.19	-22.99	-20.59	-20.59
1948 11	103.7	-5.47	-3.54	-4.50	-1.61	2.25	6.11	8.04	10.93	15.76	15.76	16.72	15.76
1953 7	136.3	-9.78	-9.05	-7.58	-6.11	-4.65	-3.18	-2.44	-.98	-.24	1.22	1.96	1.96
1957 7	145.0												

FREIGHT CARS LOADED

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	875.3	872	883	811	880	878	850	889	856	874	858	830	793
1923 5	979.0	992	985	956	934	939	964	951	944	976	984	910	886
1926 10	1036.0	1042	986	1037	1028	1040	1009	994	1001	978	984	985	966
1929 8	1030.3	1023	1009	954	959	933	943	916	943	910	912	889	857
1937 5	777.0	753	756	743	723	680	628	586	613	573	566	560	557
1945 2	839.3	852	853	834	840	837	766	768	722	782	770	784	803
1948 11	811.7	794	773	760	704	778	752	687	684	678	634	527	659
1953 7	756.0	754	734	717	692	649	676	668	647	636	661	638	658
1957 7	698.3	700	660	649	621	624	590	574	577				

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	875.3	770	720	784	776	747	708	743	745	810	742	748	782
1923 5	979.0	888	881	889	934	957	976	999	979	974	969	971	961
1926 10	1036.0	926	936	962	974	988	974	978	977	986	982	1007	1019
1929 8	1030.3	840	833	798	800	767	762	766	777	740	733	738	686
1937 5	777.0	552	574	579	585	613	617	632	622	607	613	585	603
1945 2	839.3	842	658	630	818	852	836	829	830	816	848	871	877
1948 11	811.7	682	650	629	725	726	717	761	748	790	794	795	788
1953 7	756.0	633	636	649	665	685	680	695	700	703	746	716	754
1957 7	698.3												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	875.3	-.38	.88	-7.35	.53	.30	-2.89	1.56	-2.21	-.15	-1.98	-5.18	-9.41
1923 5	979.0	1.33	.61	-2.35	-4.60	-4.09	-1.53	-2.86	-3.57	-.31	-2.55	-7.05	-9.50
1926 10	1036.0	.58	-4.83	.10	-.77	.39	-2.61	-4.05	-3.38	-5.60	-5.02	-4.92	-6.76
1929 8	1030.3	-.71	-2.07	-7.41	-6.92	-9.45	-8.48	-11.10	-8.48	-11.68	-11.48	-13.72	-16.82
1937 5	777.0	-3.09	-2.70	-4.38	-6.95	-12.48	-19.18	-24.58	-21.11	-26.25	-27.16	-27.93	-28.31
1945 2	839.3	1.51	1.63	-.64	.08	-.28	-8.74	-8.50	-13.98	-6.83	-8.26	-6.59	-4.33
1948 11	811.7	-2.18	-4.76	-6.36	-13.26	-4.15	-7.35	-15.36	-15.73	-16.47	-21.89	-35.07	-18.81
1953 7	756.0	-.26	-2.91	-5.16	-8.47	-14.15	-10.58	-11.64	-14.42	-15.87	-12.57	-15.61	-12.96
1957 7	698.3	.24	-5.49	-7.06	-11.07	-10.64	-15.51	-17.80	-17.37				

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	875.3	-12.03	-17.75	-10.43	-11.35	-14.66	-19.12	-15.12	-14.89	-7.46	-15.23	-14.55	-10.66
1923 5	979.0	-9.29	-10.01	-9.19	-4.60	-2.25	-.31	2.04	.00	-.51	-1.43	-.82	-1.84
1926 10	1036.0	-10.62	-9.65	-7.14	-5.98	-4.63	-5.98	-5.60	-5.69	-4.83	-5.21	-2.80	-1.64
1929 8	1030.3	-18.47	-19.15	-22.55	-22.35	-25.56	-26.04	-25.65	-24.59	-28.18	-28.86	-28.37	-33.42
1937 5	777.0	-28.96	-26.13	-25.48	-24.71	-21.11	-20.59	-18.66	-19.95	-21.88	-21.11	-24.71	-22.39
1945 2	839.3	.32	-21.60	-24.94	-2.54	1.51	-.40	-1.23	-1.11	-2.78	1.03	3.77	4.49
1948 11	811.7	-15.98	-19.92	-22.50	-10.68	-10.55	-11.66	-6.24	-7.84	-2.67	-2.18	-2.05	-2.92
1953 7	756.0	-16.27	-15.87	-14.15	-12.04	-9.39	-10.05	-8.07	-7.41	-7.01	-1.32	-5.29	-.26
1957 7	698.3												

BANK DEBITS OUTSIDE NEW YORK CITY

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	201.7	199	203	205	199	203	208	205	206	197	199	186	179
1923 5	193.0	193	182	186	180	187	189	185	187	197	185	191	188
1926 10	222.3	225	227	224	234	233	234	231	231	227	239	244	232
1929 8	291.0	284	298	297	249	245	242	243	241	246	239	229	227
1937 5	198.0	196	198	196	199	193	188	182	173	166	166	166	169
1945 2	445.3*	454	424	463	512	49	439	419	441	463	473	465	472
1948 11	668.3	668	650	631	647	643	635	629	628	618	632	637	629
1953 7	951.0	936	947	946	943	947	920	943	949	943	931	947	946
1937 7	1256.7	1285	1207	1195	1212	1243	1201	1197					

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	201.7	164	160	161	155	157	152	160	161	158	162	158	153
1923 5	193.0	181	187	193	190	194	196	202	212	209	206	204	206
1926 10	222.3	248	245	238	245	250	250	266	262	237	254	254	256
1929 8	291.0	222	219	205	214	207	192	188	194	191	188	182	179
1937 5	198.0	169	164	170	174	176	175	189	175	169	173	170	182
1945 2	445.3*	480	491	496	510	517	521	540	529	541	555	551	567
1948 11	668.3	650	654	658	670	676	706	725	730	772	768	765	784
1953 7	951.0	961	951	954	981	1032	1008	1019	1025	1038	1055	1061	1071
1937 7	1256.7												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	201.7	-1.32	.66	1.65	-1.32	.66	3.14	.66	2.15	-2.31	-1.32	-7.77	-14.21
1923 5	193.0	.00	-5.70	-3.63	-6.74	-3.11	-2.07	-4.15	-3.11	2.07	-4.15	-1.04	-2.59
1926 10	222.3	1.20	2.10	.75	5.25	4.80	5.25	3.90	3.90	2.10	7.50	9.74	4.35
1929 8	291.0	-2.41	2.41	2.06	-14.43	-15.81	-16.84	-16.49	-17.18	-15.46	-17.87	-21.31	-21.99
1937 5	198.0	-1.01	.00	-1.01	.50	-2.52	-5.05	-8.08	-12.63	-16.16	-16.16	-16.16	-17.68
1945 2	445.3*	1.95	-4.79	3.97	14.97	-2.54	-1.42	-5.91	-.97	3.97	6.21	4.42	5.99
1948 11	668.3	-.05	-2.74	-5.59	-3.19	-3.79	-4.99	-5.88	-6.03	-7.53	-5.44	-4.69	-5.88
1953 7	951.0	.53	-.42	-.53	-.84	-.42	-3.26	-.84	-.21	-.84	-2.10	-.42	-.53
1937 7	1256.7	2.25	-3.95	-4.91	-3.55	-1.09	-4.43	-4.75					

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	201.7	-18.68	-20.66	-20.17	-23.14	-22.15	-24.63	-20.66	-20.17	-21.65	-19.67	-21.65	-24.13
1923 5	193.0	-6.22	-3.11	.00	-1.55	.52	1.55	4.66	9.84	8.29	6.74	5.70	6.74
1926 10	222.3	11.54	10.19	7.09	10.19	12.44	12.44	19.64	17.84	6.60	14.24	14.24	15.14
1929 8	291.0	-23.71	-24.74	-29.55	-26.46	-28.87	-34.02	-35.40	-33.33	-34.36	-35.40	-37.46	-38.49
1937 5	198.0	-14.65	-17.17	-14.14	-12.12	-11.11	-11.62	-6.57	-11.62	-14.65	-11.62	-14.14	-8.08
1945 2	445.3*	7.78	10.25	11.38	14.52	16.09	16.99	21.26	17.89	21.48	24.63	23.73	27.32
1948 11	668.3	-2.74	-2.14	-1.55	.25	1.15	5.64	8.48	9.23	15.51	14.91	14.46	17.31
1953 7	951.0	1.05	.08	.31	3.15	8.52	5.99	7.15	7.78	9.15	10.94	11.57	12.62
1937 7	1256.7												

PERSONAL INCOME

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1923 5	1687.0			1686			1710			1725			1662
1926 8	1911.0			1867			1896			1918			1908
1929 8	871.7*	871	867	838	833	821	812	801	808	795	785	762	751
1937 5	759.3	768	762	764	746	733	717	698	687	687	685	677	671
1945 2	1735.7	1736	1722	1730	1751	1750	1707	1632	1668	1695	1682	1700	1693
1948 11	2128.3	2121	2087	2082	2097	2086	2081	2063	2038	2058	2071	2029	2050
1953 7	2882.0	2876	2871	2874	2865	2863	2850	2853	2847	2845	2851	2859	2859
1957 7	3462.3	3473	3472	3468	3462	3436	3436	3417	3414				

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1923 5	1687.0			1688			1765			1795			1790
1926 8	1911.0			1927			1932			1940			1984
1929 8	871.7*	745	732	716	702	693	691	728	733	684	663	652	628
1937 5	759.3	675	677	686	690	692	698	701	710	710	717	711	718
1945 2	1735.7	1726	1742	1753	1773	1819	1826	1788	1840	1839	1858	1867	1869
1948 11	2128.3	2082	2146	2175	2235	2194	2205	2222	2254	2294	2330	2352	2372
1953 7	2882.0	2872	2894	2895	2921	2935	2935	2942	2966	3009	3039	3051	3085
1957 7	3462.3												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1923 5	1687.0			-.06			1.36			2.25			-1.48
1926 8	1911.0			-2.30			-.78			.37			-.16
1929 8	871.7*	-.08	-.53	-3.86	-4.44	-5.81	-6.84	-8.11	-7.30	-8.80	-9.94	-12.58	-13.84
1937 5	759.3	1.14	.35	.61	-1.76	-3.47	-5.57	-8.08	-9.53	-9.53	-9.79	-10.84	-11.63
1945 2	1735.7	.02	-.79	-.33	.88	.83	-1.65	-5.97	-3.90	-2.34	-3.09	-2.05	-2.46
1948 11	2128.3	-.34	-1.94	-2.18	-1.47	-1.99	-2.22	-3.07	-4.24	-3.30	-2.69	-4.67	-3.68
1953 7	2882.0	-.21	-.38	-.28	-.59	-.66	-1.11	-1.01	-1.21	-1.28	-1.08	-.80	-.80
1957 7	3462.3	.31	.28	.16	-.01	-.76	-.76	-1.31	-1.40				

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1923 5	1687.0			.06			4.62			6.40			6.11
1926 8	1911.0			.84			1.10			1.52			3.82
1929 8	871.7*	-14.53	-16.02	-17.86	-19.46	-20.50	-20.73	-16.48	-15.91	-21.53	-23.94	-25.20	-27.95
1937 5	759.3	-11.11	-10.84	-9.66	-9.13	-8.87	-8.08	-7.68	-6.50	-6.50	-5.57	-6.36	-5.44
1945 2	1735.7	-.56	.36	1.00	2.15	4.80	5.20	3.02	6.01	5.95	7.05	7.57	7.88
1948 11	2128.3	-2.18	.83	2.19	5.01	3.09	3.60	4.40	5.90	7.78	9.48	10.51	11.45
1953 7	2882.0	-.35	.42	.45	1.35	1.84	1.84	2.08	2.91	4.41	5.45	5.86	7.04
1957 7	3462.3												

RETAIL STORE SALES

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	31.0	31	32	32	34	33	34	33	33	31	33	31	32
1923 5	34.3	35	34	35	34	35	34	34	35	35	34	34	34
1926 10	37.0	37	37	37	37	37	36	36	36	37	38	36	37
1929 8	38.3	39	38	38	38	37	37	37	36	36	35	35	35
1937 5	3583.3*	3553	3601	3630	3648	3593	3346	3254	3262	3174	3119	3102	3039
1945 2	6256.0	6346	6109	6253	6404	6572	6467	6497	6835	7065	7064	7522	7692
1948 11	11014.7	11146	10694	10846	10888	11060	10965	10940	10742	10855	11066	10900	10964
1953 7	14184.3	13939	13834	13911	13988	13806	13512	13836	13792	14146	13905	14292	14144
1957 7	16949.0	17030	16919	16714	16562	16855	16718	16115	15928				

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	31.0	32	31	31	30	30	30	29	28	29	29	29	29
1923 5	34.3	34	34	33	34	33	35	34	34	35	35	35	35
1926 10	37.0	37	37	37	37	36	36	37	37	38	37	39	37
1929 8	38.3	35	35	34	33	34	34	34	35	33	33	32	31
1937 5	3583.3*	3078	3116	3165	3192	3180	3282	3344	3390	3381	3419	3416	3452
1945 2	6256.0	7797	8076	8334	8256	8679	9168	9040	9136	9341	9447	9367	9572
1948 11	11014.7	10801	11075	11309	11395	11432	11656	12069	12985	12965	12381	12043	11768
1953 7	14184.3	14006	14070	13946	14233	14995	14736	14631	14936	15131	15232	15207	15148
1957 7	16949.0												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	31.0	.00	3.23	3.23	9.68	6.45	9.68	6.45	6.45	.00	6.45	.00	3.23
1923 5	34.3	1.94	-.97	1.94	-.97	1.94	-.97	-.97	1.94	1.94	-.97	-.97	-.97
1926 10	37.0	.00	.00	.00	.00	.00	-2.70	-2.70	-2.70	.00	2.70	-2.70	.00
1929 8	38.3	1.74	-.87	-.87	-.87	-3.48	-3.48	-3.48	-6.09	-6.09	-8.70	-8.70	-8.70
1937 5	3583.3*	-.85	.49	1.30	1.80	.27	-6.62	-9.19	-8.97	-11.42	-12.96	-13.43	-15.19
1945 2	6256.0	1.44	-2.35	-.05	2.37	5.05	3.37	3.85	9.25	12.93	12.92	20.24	22.95
1948 11	11014.7	1.19	-2.91	-1.53	-1.15	.41	-.45	-.68	-2.48	-1.45	.47	-1.04	-.46
1953 7	14184.3	-1.73	-2.47	-1.93	-1.38	-2.67	-4.74	-2.46	-2.77	-.27	-1.97	.76	-.28
1957 7	16949.0	.48	-.18	-1.39	-2.28	-.55	-1.36	-4.92	-6.02				

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	31.0	3.23	.00	.00	-3.23	-3.23	-3.23	-6.45	-9.68	-6.45	-6.45	-6.45	-6.45
1923 5	34.3	-.97	-.97	-3.88	-.97	-3.88	1.94	-.97	-.97	1.94	1.94	1.94	1.94
1926 10	37.0	.00	.00	.00	.00	-2.70	-2.70	.00	.00	2.70	.00	5.40	.00
1929 8	38.3	-8.70	-8.70	-11.30	-13.91	-11.30	-11.30	-11.30	-8.70	-13.91	-13.91	-16.52	-19.13
1937 5	3583.3*	-14.10	-13.04	-11.67	-10.92	-11.26	-8.41	-6.68	-5.39	-5.65	-4.59	-4.67	-3.66
1945 2	6256.0	24.63	29.09	33.22	31.97	38.73	46.55	44.50	46.04	49.31	51.01	49.73	53.00
1948 11	11014.7	-1.94	.55	2.67	3.45	3.79	5.82	9.57	17.89	17.71	12.40	9.34	6.84
1953 7	14184.3	-1.26	-.81	-1.68	.34	5.72	3.89	3.15	5.30	6.67	7.39	7.21	8.20
1957 7	16949.0												

WHOLESALE PRICES EX FARM AND FOODS

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	1526.0	1592	1628	1690	1706	1705	1794	1742	1674	1580	1440	1333	1242
1923 5	1075.0	1051	1034	1020	1018	1006	993	995	1024	1035	1028	1014	1000
1926 10	710.7*	709	701	690	685	676	669	668	668	668	670	671	669
1929 8	654.3	655	655	649	647	640	636	631	628	624	613	604	598
1937 5	616.7	615	617	616	614	609	602	598	597	594	591	586	584
1945 2	709.0	710	710	711	712	713	714	714	715	717	719	721	724
1948 11	1053.7	1054	1049	1040	1033	1020	1008	1001	998	1000	1000	1000	1000
1953 7	1145.3	1149	1147	1146	1145	1146	1146	1144	1142	1145	1145	1142	1143
1957 7	1256.3	1260	1260	1258	1259	1261	1262	1258	1259				

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	1526.0	1141	1100	1083	1055	1024	992	967	968	998	1014	1001	981
1923 5	1075.0	981	974	977	977	977	988	1006	1014	1040	1028	1013	1016
1926 10	710.7*	665	667	664	664	662	663	664	662	662	664	665	664
1929 8	654.3	595	587	580	574	565	560	552	542	537	530	529	530
1937 5	616.7	581	582	582	582	580	576	574	573	574	575	576	576
1945 2	709.0	731	738	743	755	783	798	802	828	863	892	918	922
1948 11	1053.7	1002	1004	1006	1007	1008	1015	1022	1038	1059	1082	1101	1116
1953 7	1145.3	1144	1144	1145	1148	1149	1152	1157	1156	1157	1155	1156	1165
1957 7	1256.3												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	1526.0	4.32	6.68	10.75	11.80	11.73	13.63	14.15	9.70	3.54	-5.64	-12.65	-18.61
1923 5	1075.0	-2.23	-3.81	-5.12	-5.30	-6.42	-7.63	-7.44	-4.74	-3.72	-4.37	-5.67	-6.98
1926 10	710.7*	-.23	-1.36	-2.91	-3.61	-4.88	-5.86	-6.00	-6.00	-6.00	-5.72	-5.58	-5.86
1929 8	654.3	.10	-.81	-1.12	-2.19	-2.80	-3.57	-4.02	-4.64	-6.32	-7.69	-8.61	
1937 5	616.7	-.27	.05	-.11	-.43	-1.24	-2.38	-3.03	-3.19	-3.68	-4.16	-4.97	-5.30
1945 2	709.0	.14	.14	.28	.42	.56	.70	.70	.85	1.13	1.41	1.69	2.12
1948 11	1053.7	.03	-.44	-1.30	-1.96	-3.20	-4.33	-5.00	-5.28	-5.09	-5.09	-5.09	-5.09
1953 7	1145.3	.32	.15	.06	-.03	.06	.06	-.12	-.29	-.03	-.03	-.29	-.20
1957 7	1256.3	.29	.29	.13	.21	.37	.45	.13	.21				

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	1526.0	-25.23	-27.92	-29.03	-30.86	-32.90	-34.99	-36.63	-36.57	-34.60	-33.55	-34.40	-35.71
1923 5	1075.0	-8.74	-9.40	-9.12	-9.12	-9.12	-8.09	-6.42	-5.67	-3.26	-4.37	-5.77	-5.49
1926 10	710.7*	-6.43	-6.14	-6.57	-6.57	-6.85	-6.71	-6.57	-6.85	-6.85	-6.57	-6.43	-6.57
1929 8	654.3	-9.07	-10.29	-11.36	-12.28	-13.65	-14.42	-15.64	-17.17	-17.93	-19.00	-19.15	-19.00
1937 5	616.7	-5.78	-5.62	-5.62	-5.62	-5.95	-6.59	-6.92	-7.08	-6.92	-6.76	-6.59	-6.59
1945 2	709.0	3.10	4.09	4.80	6.49	10.44	12.55	13.12	16.78	21.72	25.81	29.48	30.04
1948 11	1053.7	-4.90	-4.71	-4.52	-4.43	-4.33	-3.67	-3.01	-1.49	.51	2.69	4.49	5.92
1953 7	1145.3	-.12	-.12	-.03	.23	.32	.58	1.02	.93	1.02	.84	.93	1.72
1957 7	1256.3												

CORPORATE PROFITS AFTER TAXES

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 3 MO	+ 6 MO	+ 9 MO	+12 MO
1920 2	498.0	446	328	52	46
1923 5	516.0	376	310	434	294
1926 8	530.0	497	452	394	411
1929 8	678.0	562	394	291	149
1937 5	427.0	360	204	118	79
1945 2	102.0*	96	69	65	78
1948 11	195.0	169	162	152	151
1953 5	182.0	175	134	154	160
1957 8	204.0				

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+15 MO	+18 MO	+21 MO	+24 MO
1920 2	498.0				157
1923 5	516.0	276	392	448	469
1926 8	530.0	391*	442	477	559
1929 8	678.0	53	34		
1937 5	427.0	143	282		
1945 2	102.0*	113	158	183	189
1948 11	195.0	157	206	251	256
1953 5	182.0	157	170	195	201
1957 8	204.0				

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 3 MO	+ 6 MO	+ 9 MO	+12 MO
1920 2	498.0	-10.44	-34.14	-89.56	-90.76
1923 5	516.0	-27.13	-39.92	-15.89	-43.02
1926 8	530.0	-6.23	-14.72	-25.66	-22.45
1929 8	678.0	-17.11	-41.89	-57.08	-78.91
1937 5	427.0	-15.69	-52.22	-72.36	-81.50
1945 2	102.0*	-5.88	-32.35	-36.27	-23.53
1948 11	195.0	-13.33	-16.92	-22.05	-22.56
1953 5	182.0	-3.85	-26.37	-15.38	-12.09
1957 8	204.0				

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+15 MO	+18 MO	+21 MO	+24 MO
1920 2	498.0				-68.47
1923 5	516.0	-46.51	-24.03	-13.18	-9.11
1926 8	530.0	-26.23	-16.60	-10.00	5.47
1929 8	678.0	-92.18	-94.99		
1937 5	427.0	-66.51	-33.96		
1945 2	102.0*	10.78	54.90	79.41	85.29
1948 11	195.0	-19.49	5.64	28.72	31.28
1953 5	182.0	-13.74	-6.59	7.14	10.44
1957 8	204.0				

AVERAGE WORK WEEK IN MANUFACTURING

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1923 5	498.0	496	498	493	480	484	488	482	479	476	471	474	466
1926 10	484.3	482	477	478	478	478	476	482	478	476	481	477	479
1929 8	488.7	488	492	468	466	459	461	454	455	449	446	435	428
1937 5	400.7*	393	387	383	373	368	355	342	343	348	347	347	347
1945 2	452.7	452	451	442	446	445	408	417	414	411	412	408	405
1948 11	396.7	396	395	394	391	388	389	388	392	391	395	394	390
1953 7	406.0	405	399	401	398	397	395	396	394	393	395	395	397
1957 7	399.7	400	398	393	391	389	388	384	385				

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1923 5	498.0	458	456	460	464	474	475	480	482	479	481	478	480
1926 10	484.3	477	473	477	478	477	471	477	481	482	484	483	478
1929 8	488.7	424	425	415	414	400	418	418	417	428	414	406	401
1937 5	400.7*	346	357	362	369	368	366	371	376	376	375	371	370
1945 2	452.7	407	409	401	401	401	405	401	402	402	403	405	404
1948 11	396.7	393	397	397	397	401	402	405	409	412	409	410	410
1953 7	406.0	397	397	397	400	400	403	404	406	406	410	407	408
1957 7	399.7												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1923 5	498.0	-.40	.00	-1.00	-3.61	-2.81	-2.01	-3.21	-3.81	-4.42	-5.42	-4.82	-6.43
1926 10	484.3	-.48	-1.51	-1.31	-1.31	-1.31	-1.72	-.48	-1.31	-1.72	-.69	-1.51	-2.34
1929 8	488.7	-.14	.68	-4.23	-4.64	-6.07	-5.66	-7.09	-6.89	-8.12	-8.73	-10.98	-12.41
1937 5	400.7*	-1.91	-3.41	-4.41	-6.90	-8.15	-11.40	-14.64	-14.39	-13.14	-13.39	-13.39	-13.39
1945 2	452.7	-.15	-.37	-2.36	-1.47	-1.69	-9.87	-7.88	-8.54	-9.20	-8.98	-9.87	-10.53
1948 11	396.7	-.17	-.42	-.67	-1.43	-2.18	-1.93	-2.18	-1.18	-1.43	-.42	-.67	-1.68
1953 7	406.0	-.25	-1.72	-1.23	-1.97	-2.22	-2.71	-2.46	-2.96	-3.20	-2.71	-2.71	-2.22
1957 7	399.7	.08	-.42	-1.67	-2.17	-2.67	-2.92	-3.92	-3.67				

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1923 5	498.0	-8.03	-8.43	-7.63	-6.83	-4.82	-4.62	-3.61	-3.21	-3.81	-3.41	-4.02	-3.61
1926 10	484.3	-1.51	-2.34	-1.51	-1.31	-1.51	-2.75	-1.51	-.69	-.48	-.07	-.27	-1.31
1929 8	488.7	-13.23	-13.03	-15.07	-15.28	-18.14	-14.46	-14.46	-14.67	-12.41	-15.28	-16.92	-17.94
1937 5	400.7*	-13.64	-10.90	-9.65	-7.90	-8.15	-8.65	-7.40	-6.16	-6.16	-6.41	-7.40	-7.65
1945 2	452.7	-10.09	-9.65	-11.41	-11.41	-11.41	-10.53	-11.41	-11.19	-11.19	-10.97	-10.53	-10.75
1948 11	396.7	-.92	.08	.08	.08	1.09	1.34	2.10	3.11	3.87	3.11	3.36	3.36
1953 7	406.0	-2.22	-2.22	-2.22	-1.48	-1.48	-.74	-.49	.00	.00	.98	.25	.00
1957 7	399.7												

LAYOFF RATE MANUFACTURING

DATE	AT PEAK	STANDINGS + 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	FIRST YEAR AFTER PEAK + 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	6.7	6	7	1	5	8	22	32	21	131	302	423	195
1923 5	20.7	23	34	34	42	40	64	31	46	31	45	96	78
1926 10	50.3	59	61	102	82	52	72	57	64	54	84	66	75
1929 8	52.3	65	83	106	122	101	118	99	105	104	109	139	151
1937 5	177.7*	388	323	274	308	255	237	243	193	193	230	260	270
1945 2	66.7	150	120	120	110	70	80	110	100	80	100	90	90
1948 11	163.3	200	150	170	140	120	110	100	60	60	80	80	100
1953 7	113.3	170	180	170	150	150	140	110	130	120	110	130	130
1957 7	136.7												

DATE	AT PEAK	STANDINGS +13 MO	+14 MO	+15 MO	+16 MO	+17 MO	SECOND YEAR AFTER PEAK +18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	6.7	216	259	271	207	260	223	182	229	104	55	102	38
1923 5	20.7	83	58	59	72	67	66	34	29	30	45	52	41
1926 10	50.3	64	74	70	70	68	52	62	60	53	50	51	42
1929 8	52.3	148	150	102	94	97	85	74	87	95	131	118	93
1937 5	177.7*	204	212	302	334	473	582	589	470	391	386	385	386
1945 2	66.7	60	70	100	160	180	1290	510	230	190	130	170	180
1948 11	163.3	220	240	250	240	240	280	240	250	220	200	230	270
1953 7	113.3	130	160	190	220	220	250	220	240	240	190	190	160
1957 7	136.7	160	200	240	250	230	350	260					

DATE	AT PEAK	ABSOLUTE CHANGES + 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	FIRST YEAR AFTER PEAK + 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	6.7	-.7	.3	-5.7	-1.7	1.3	15.3	25.3	14.3	124.3	295.3	416.3	188.3
1923 5	20.7	2.3	13.3	13.3	21.3	19.3	43.3	10.3	25.3	10.3	24.3	35.3	57.3
1926 10	50.3	8.7	10.7	51.7	31.7	1.7	21.7	6.7	13.7	3.7	33.7	15.7	24.7
1929 8	52.3	12.7	30.7	53.7	69.7	48.7	65.7	46.7	52.7	51.7	56.7	85.7	98.7
1937 5	177.7*	26.3	34.3	124.3	156.3	295.3	404.3	411.3	292.3	213.3	208.3	207.3	208.3
1945 2	66.7	-6.7	3.3	33.3	93.3	113.3	1223.3	443.3	163.3	123.3	63.3	103.3	113.3
1948 11	163.3	56.7	76.7	86.7	76.7	76.7	116.7	76.7	86.7	56.7	36.7	66.7	106.7
1953 7	113.3	16.7	46.7	76.7	106.7	106.7	136.7	106.7	126.7	126.7	76.7	76.7	46.7
1957 7	136.7	23.3	63.3	103.3	113.3	93.3	213.3	123.3					

DATE	AT PEAK	ABSOLUTE CHANGES +13 MO	+14 MO	+15 MO	+16 MO	+17 MO	SECOND YEAR AFTER PEAK +18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	6.7	209.3	252.3	264.3	200.3	253.3	216.3	175.3	222.3	97.3	48.3	95.3	31.3
1923 5	20.7	62.3	37.3	38.3	51.3	46.3	45.3	13.3	8.3	9.3	24.3	31.3	20.3
1926 10	50.3	13.7	23.7	19.7	19.7	17.7	1.7	11.7	9.7	2.7	-.3	.7	-8.3
1929 8	52.3	95.7	97.7	49.7	41.7	44.7	32.7	21.7	34.7	42.7	78.7	65.7	40.7
1937 5	177.7*	210.3	145.3	96.3	130.3	77.3	59.3	65.3	15.3	15.3	52.3	82.3	92.3
1945 2	66.7	83.3	53.3	53.3	43.3	3.3	13.3	43.3	33.3	13.3	33.3	23.3	23.3
1948 11	163.3	36.7	-13.3	6.7	-23.3	-43.3	-53.3	-63.3	-103.3	-103.3	-83.3	-83.3	-63.3
1953 7	113.3	56.7	66.7	56.7	36.7	36.7	26.7	-3.3	16.7	6.7	-3.3	16.7	16.7
1957 7	136.7												

GROSS ACCESSION RATE MANUFG

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	1190.0	1266	1278	1010	975	1352	1253	1237	1045	592	426	342	299
1923 5	1105.7	1156	961	721	787	575	469	478	556	478	361	371	233
1926 10	429.0	393	371	331	372	374	379	378	352	306	282	289	307
1929 8	426.7	396	346	232	200	216	240	199	194	194	154	141	105
1937 5	388.0*	342	311	275	326	261	195	262	367	352	326	280	305
1945 2	576.7	520	510	540	540	560	560	670	780	880	830	790	760
1948 11	406.7	370	330	330	320	320	350	370	340	360	360	340	370
1953 7	400.0	380	340	290	300	300	290	280	290	260	270	280	290
1957 7	303.3	280	280	260	240	240	260	240					

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	1190.0	280	292	274	288	160	101	149	200	435	454	483	560
1923 5	1105.7	236	216	260	292	239	426	472	484	442	447	450	358
1926 10	429.0	310	324	305	291	311	319	371	323	379	397	377	429
1929 8	426.7	147	133	100	113	132	155	199	167	155	126	151	128
1937 5	388.0*	319	445	434	389	476	461	398	397	344	348	318	354
1945 2	576.7	750	730	660	610	720	670	640	620	580	520	560	560
1948 11	406.7	440	370	360	380	390	440	400	460	540	500	480	440
1953 7	400.0	290	290	320	360	360	350	360	380	370	380	350	340
1957 7	303.3												

ABSOLUTE CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	1190.0	76.0	88.0	-180.0	-215.0	162.0	63.0	47.0	-145.0	-598.0	-764.0	-848.0	-891.0
1923 5	1105.7	50.3	-144.7	-384.7	-318.7	-530.7	-636.7	-627.7	-549.7	-627.7	-744.7	-734.7	-872.7
1926 10	429.0	-36.0	-58.0	-98.0	-57.0	-55.0	-50.0	-51.0	-77.0	-123.0	-147.0	-140.0	-122.0
1929 8	426.7	-30.7	-80.7	-194.7	-226.7	-210.7	-186.7	-227.7	-232.7	-232.7	-272.7	-285.7	-321.7
1937 5	388.0*	-46.0	-77.0	-113.0	-62.0	-127.0	-193.0	-126.0	-21.0	-36.0	-62.0	-108.0	-83.0
1945 2	576.7	-56.7	-66.7	-36.7	-36.7	-16.7	-16.7	93.3	203.3	303.3	253.3	213.3	183.3
1948 11	406.7	-36.7	-76.7	-76.7	-86.7	-86.7	-56.7	-36.7	-66.7	-46.7	-46.7	-66.7	-36.7
1953 7	400.0	-20.0	-60.0	-110.0	-100.0	-100.0	-110.0	-120.0	-110.0	-140.0	-130.0	-120.0	-110.0
1957 7	303.3	-23.3	-23.3	-43.3	-63.3	-63.3	-43.3	-63.3					

ABSOLUTE CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	1190.0	-910.0	-898.0	-916.0	-902.0	-1030.0	-1089.0	-1041.0	-990.0	-755.0	-736.0	-707.0	-630.0
1923 5	1105.7	-869.7	-889.7	-845.7	-813.7	-866.7	-679.7	-633.7	-621.7	-663.7	-658.7	-655.7	-747.7
1926 10	429.0	-119.0	-105.0	-124.0	-138.0	-118.0	-110.0	-58.0	-106.0	-50.0	-92.0	-52.0	.0
1929 8	426.7	-279.7	-293.7	-326.7	-313.7	-294.7	-271.7	-227.7	-259.7	-271.7	-300.7	-275.7	-298.7
1937 5	388.0*	-69.0	57.0	46.0	1.0	88.0	73.0	10.0	9.0	-44.0	-40.0	-70.0	-34.0
1945 2	576.7	173.3	153.3	83.3	33.3	143.3	93.3	63.3	43.3	3.3	-56.7	-16.7	-16.7
1948 11	406.7	33.3	-36.7	-46.7	-26.7	-16.7	33.3	-6.7	53.3	133.3	93.3	73.3	33.3
1953 7	400.0	-110.0	-110.0	-80.0	-40.0	-40.0	-50.0	-40.0	-20.0	-30.0	-20.0	-50.0	-60.0
1957 7	303.3												

NEW ORDERS FOR DURABLE GOODS

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	564.3	586	458	610	499	470	516	439	372	327	312	286	248
1923 5	857.3	900	802	851	920	886	859	1093	1082	981	973	762	741
1926 10	974.7	988	872	1070	1070	980	940	1030	930	870	920	900	870
1929 8	151.0*	145	139	114	127	128	132	108	111	106	107	97	91
1937 5	122.0	107	115	104	98	92	84	76	74	77	79	70	71
1945 2	8158.0*	7040	6673	4776	2767	4547	*	1936	4040	4838	5127	5834	4508
1948 11	7485.3	7466	6329	6911	6735	6223	6145	6241	5870	6669	6992	6621	7151
1953 7	11125.3	10023	9638	9622	9416	8928	9097	9842	9534	9446	9510	9743	9823
1957 7	13138.0	13160	12519	12154	12362	11399	10704	10549					

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	564.3	285	289	362	362	382	354	459	501	614	581	612	551
1923 5	857.3	847	1021	1013	1086	1096	1235	1148	1077	925	882	928	924
1926 10	974.7	970	1080	1000	1010	960	980	1030	990	980	970	910	900
1929 8	151.0*	93	81	74	73	61	69	73	67	72	63	58	57
1937 5	122.0	73	76	83	82	96	99	98	96	98	98	89	102
1945 2	8158.0*	5510	6399	7176	5615	6188	6026	5717	6158	5916	6315	5998	6134
1948 11	7485.3	7041	7468	7514	8081	8198	9434	10083	11437	14260	12037	11852	11023
1953 7	11125.3	10262	10942	11103	10856	11784	12404	12514	13538	12489	13562	13445	13760
1957 7	13138.0												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	564.3	3.84	-18.84	8.09	-11.58	-16.72	-8.56	-22.21	-34.08	-42.06	-44.71	-52.86	-56.05
1923 5	857.3	4.98	-6.45	-.74	7.31	3.34	-.50	27.49	26.20	14.42	13.49	-11.12	-13.57
1926 10	974.7	1.37	-10.53	9.78	9.78	.55	-3.56	5.68	-4.58	-10.74	-5.61	-7.66	-10.74
1929 8	151.0*	-3.97	-7.95	-24.50	-15.89	-15.23	-12.58	-29.80	-26.49	-29.80	-29.14	-35.76	-39.73
1937 5	122.0	-12.29	-5.74	-14.75	-19.67	-24.59	-31.15	-37.70	-39.34	-36.89	-35.25	-42.62	-41.80
1945 2	8158.0*	-13.70	-18.20	-41.46	-66.08	-44.26	-99.99	-76.27	-50.48	-40.70	-37.15	-28.49	-44.74
1948 11	7485.3	-.26	-15.45	-7.67	-10.02	-16.86	-17.91	-16.62	-21.58	-10.91	-6.59	-11.55	-4.47
1953 7	11125.3	-9.91	-13.37	-13.51	-15.36	-19.75	-18.23	-11.53	-14.30	-15.09	-14.52	-12.42	-11.71
1957 7	13138.0	.17	-4.71	-7.49	-5.91	-13.24	-18.53	-19.71					

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	564.3	-49.50	-48.79	-35.85	-35.85	-32.31	-37.27	-18.66	-11.22	8.80	2.95	-26.99	-2.36
1923 5	857.3	-1.20	19.09	18.16	26.67	27.84	43.82	33.90	25.62	7.89	2.88	8.24	7.78
1926 10	974.7	-.48	10.81	2.60	3.62	-1.50	.55	5.68	1.57	.55	-.48	-6.63	-7.66
1929 8	151.0*	-38.41	-46.36	-50.99	-51.66	-59.60	-54.30	-51.66	-55.63	-52.32	-58.28	-61.59	-62.25
1937 5	122.0	-40.16	-37.70	-31.97	-32.79	-21.31	-18.85	-19.67	-21.31	-19.67	-19.67	-27.05	-16.39
1945 2	8158.0*	-32.46	-21.56	-12.04	-31.17	-24.15	-26.13	-29.92	-24.52	-27.48	-22.59	-26.48	-24.56
1948 11	7485.3	-5.94	-.23	.38	7.96	9.52	26.03	34.70	52.79	90.51	60.81	58.34	47.26
1953 7	11125.3	-7.76	-1.65	-.20	-2.42	5.92	11.49	12.48	21.69	12.26	21.90	20.85	23.68
1957 7	13138.0												

RESIDENTIAL CONTRACTS

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	1805.3	1559	1564	1711	1336	844	861	964	798	980	925	768	830
1923 5	3121.0*	2709	3120	2954	2750	3817	3511	3492	4677	4173	4024	3727	3369
1926 10	4316.0*	4187	4019	4311	4048	4325	4003	3440	4080	3773	4248	4011	4146
1929 8	3186.3	2581	2528	2182	2194	1895	1901	1785	2048	2022	1942	1899	1736
1937 5	2065.7	1890	1888	1892	1738	1443	1630	1375	1376	1522	1911	1801	1756
1945 2	504.0	512	881	833	757	855	769	944	1088	1607	1783	2321	2392
1948 11	3531.3	3750	2617	3332	2915	3472	3647	4633	4102	4136	5884	5477	5548
1953 7	5589.3	5324	5582	6338	5710	5754	6174	6129	6273	6538	6796	7171	6987
1957 7	10217.0*	10551	10667	10534	9930	9283	9564	7840					

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	1805.3	1143	1169	1403	1589	1749	1623	1930	2357	2132	2201	2305	2231
1923 5	3121.0*	3271	2962	3011	3070	3120	3356	3489	3305	3713	3702	4020	4076
1926 10	4316.0*	4382	4937	5358	5524	4899	4596	5054	5021	4653	4596	4114	4349
1929 8	3186.3	1989	1960	1922	1815	1677	2070	1904	1825	1889	1580	1655	1456
1937 5	2065.7	1970	2095	2104	1998	2471	2187	2553	2938	2821	2926	2729	2786
1945 2	504.0	4731	5371	6097	5052	4626	4560	5178	3418	3387	3123	5168	4059
1948 11	3531.3	5890	5537	5780	7375	7867	7198	7785	7955	7545	6081	5850	6334
1953 7	5589.3	7059	8404	8251	8157	9543	8954	8667	8902	8768	7780	9361	8831
1957 7	10217.0*												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	1805.3	-13.64	-13.37	-5.22	-26.00	-53.25	-52.31	-46.60	-55.80	-45.72	-48.76	-57.46	-54.02
1923 5	3121.0*	-13.20	-.03	-5.35	-11.89	22.30	12.50	11.89	49.86	33.71	28.93	19.42	7.95
1926 10	4316.0*	-2.99	-6.88	-.12	-6.21	.21	-7.25	-20.30	-5.47	-12.58	-1.58	-7.07	-3.94
1929 8	3186.3	-19.00	-20.66	-31.52	-31.14	-40.53	-40.34	-43.98	-35.73	-36.54	-39.05	-40.40	-45.52
1937 5	2065.7	-8.50	-8.60	-8.41	-15.86	-30.14	-21.09	-33.44	-33.39	-26.32	-7.49	-12.81	-14.99
1945 2	504.0	1.59	74.80	65.28	50.20	69.64	52.58	87.30	115.87	218.85	253.77	360.52	374.60
1948 11	3531.3	6.19	-25.89	-5.64	-17.45	-1.68	3.28	31.20	16.16	17.12	66.62	55.10	57.11
1953 7	5589.3	-4.75	-.13	13.39	2.16	2.95	10.46	9.65	12.23	16.97	21.59	28.30	25.01
1957 7	10217.0*	3.27	4.40	3.10	-2.81	-9.14	-6.39	-23.26					

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	1805.3	-36.69	-35.25	-22.29	-11.98	-3.12	-10.10	6.91	29.45	18.09	21.92	27.68	23.63
1923 5	3121.0*	4.81	-5.09	-3.52	-1.63	-.09	7.53	11.79	5.90	18.97	18.62	28.80	30.60
1926 10	4316.0*	1.53	14.39	24.14	27.99	13.51	6.49	17.10	16.33	7.81	6.49	-4.68	.76
1929 8	3186.3	-37.58	-38.49	-39.68	-43.04	-47.37	-35.03	-40.24	-42.72	-40.72	-50.41	-48.06	-54.30
1937 5	2065.7	-4.63	1.42	1.86	-3.28	19.62	5.87	23.59	42.23	36.57	41.65	32.11	34.87
1945 2	504.0	838.69	965.67	1109.72	902.38	817.86	804.76	927.38	578.17	572.02	519.64	925.40	705.36
1948 11	3531.3	66.79	56.80	63.68	108.84	122.78	103.83	120.45	125.27	113.66	72.20	65.66	79.37
1953 7	5589.3	26.29	50.36	47.62	45.94	70.74	60.20	55.06	59.27	56.87	39.19	67.48	58.00
1957 7	10217.0*												

COMMERCIAL AND INDUSTRIAL CONTRACTS

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	3324.0	2877	2842	2072	1726	1890	1434	1211	1442	975	1020	767	793
1923 5	1423.7	1252	1092	921	1107	1124	1278	1350	1139	1495	1213	1072	1143
1926 10	1914.0*	1762	1811	1742	1786	1931	1926	1884	1830	1715	1688	1572	1767
1929 8	2256.0*	2073	2097	2318	2345	1360	1730	1643	1517	1454	1226	1212	1106
1937 5	1152.7	1217	1343	1288	1012	723	703	735	532	539	633	580	616
1945 2	1293.3	1755	1340	979	960	1446	1415	2806	2653	3512	2892	3714	3670
1948 11	1488.7	1313	1350	1258	1412	1376	1188	1269	1165	902	1361	1214	1280
1953 7	2040.3	1860	1966	2230	2200	1301	1924	1616	1800	1935	2201	2138	1933
1957 7	3661.0*	3614	3323	3041	3377	2593	3259	2861					

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	3324.0	714	617	832	852	872	909	937	1004	1161	1057	943	984
1923 5	1423.7	1144	878	943	1072	1222	1338	1273	1582	997	1271	1511	1549
1926 10	1914.0*	1574	1626										
1929 8	2256.0*	1091	870	795	789	772	672	618	675	634	654	616	418
1937 5	1152.7	567	524	608	468	672	620	593	753	760	763	750	746
1945 2	1293.3	3955	3379	3911	3811	3052	2359	2318	2628	1909	1451	2368	1992
1948 11	1488.7	1299	1288	1242	1712	1942	2021	2001	1973	2475	2230	2266	2456
1953 7	2040.3	2025	2031	2200	1838	2061	2336	2615	2582	2183	2316	2488	2815
1957 7	3661.0*												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	3324.0	-13.45	-14.50	-37.67	-48.07	-43.14	-56.86	-63.57	-56.62	-70.67	-69.31	-76.93	-77.95
1923 5	1423.7	-12.06	-23.30	-35.31	-22.24	-21.05	-10.23	-5.17	-20.00	5.01	-14.80	-24.70	-19.71
1926 10	1914.0*	-7.94	-5.38	-8.99	-6.69	.89	.63	-1.57	-4.39	-10.40	-11.81	-17.87	-7.68
1929 8	2256.0*	-8.11	-7.05	2.75	3.94	-39.72	-23.32	-27.17	-32.76	-35.55	-45.66	-46.28	-50.97
1937 5	1152.7	5.58	16.51	11.74	-12.20	-37.28	-39.01	-36.23	-53.85	-53.24	-45.08	-49.68	-46.56
1945 2	1293.3	35.70	3.61	-24.30	-25.77	11.80	9.41	116.96	105.13	171.55	123.61	187.16	183.76
1948 11	1488.7	-11.80	-9.31	-15.49	-5.15	-7.57	-20.20	-14.76	-21.74	-39.41	-8.58	-18.45	-14.02
1953 7	2040.3	-8.84	-3.64	9.30	7.83	-36.24	-5.70	-20.80	-11.78	-5.16	7.87	4.79	-5.26
1957 7	3661.0*	-1.28	-9.23	-16.93	-7.76	-29.17	-10.98	-21.85					

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	3324.0	-78.52	-81.44	-74.97	-74.37	-73.77	-72.65	-71.81	-69.80	-65.07	-68.20	-71.63	-70.40
1923 5	1423.7	-19.64	-38.33	-33.76	-24.70	-14.16	-6.02	-10.58	11.12	-29.97	-10.72	6.13	8.80
1926 10	1914.0*	-17.76	-15.05										
1929 8	2256.0*	-51.64	-61.44	-64.76	-65.03	-65.78	-70.21	-72.61	-70.08	-71.90	-71.01	-72.69	-81.47
1937 5	1152.7	-50.81	-54.54	-47.25	-59.40	-41.70	-46.21	-48.55	-34.67	-34.07	-33.81	-34.93	-35.28
1945 2	1293.3	205.80	161.26	202.40	194.66	135.98	82.40	79.23	103.20	47.60	12.19	83.09	49.38
1948 11	1488.7	-12.74	-13.48	-16.57	15.00	30.45	35.76	34.42	32.53	66.26	49.80	52.22	64.98
1953 7	2040.3	-.75	-.46	7.83	-9.92	1.01	14.49	28.16	26.55	6.99	13.51	21.94	37.97
1957 7	3661.0*												

NUMBER OF NEW INCORPORATIONS

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	1312.7	1282	1268	1250	1164	1192	1138	1101	1109	1105	1051	913	893
1923 5	1053.3	1017	1037	1036	857	915	1034	1004	964	965	922	914	878
1926 10	3567.3*	3549	3442	3545	3515	3770	3707	3548	3681	3542	3675	3748	3568
1929 8	3911.0	3797	3848	3747	3523	3529	3628	3521	3543	3716	3483	3574	3224
1937 5	2301.7*	2279	2161	2110	2078	2002	2070	2043	1939	1933	1976	1913	1981
1945 2	1397.7*	1386	1488	1614	1659	1717	1975	2409	2980	3202	3405	4347	4805
1948 11	7093.0*	7136	6935	6697	6758	6612	6958	6981	6692	7422	7716	7239	7855
1953 7	8595.7	8138	8352	8702	8452	9005	8371	8982	9304	9338	8673	9373	9801
1957 7	11786.0	12349	11827	11843	10779	10682	11474	11017					

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	1312.7	922	977	966	932	935	914	1001	970	904	984	911	925
1923 5	1053.3	879	890	842	971	990	875	1038					
1926 10	3567.3*	3821	3719	3723	3990	3789	3755	4147	4055	3849	3899	3630	4101
1929 8	3911.0	3614	3586	3483	3584	3358	3417	3591	3547	3446	3585	3606	3406
1937 5	2301.7*	1977	1988	2086	1990	1897	2000	1969	1964	1976	2091	1909	2248
1945 2	1397.7*	4462	4612	4588	4302	4289	4080	3819	3674	3371	3359	3309	3316
1948 11	7093.0*	7936	7956	8143	8124	7614	8613	8520	7491	7827	7053	7139	7274
1953 7	8595.7	9827	10400	10371	11320	12102	11562	11967	11873	10687	11242	12120	11347
1957 7	11786.0												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	1312.7	-2.34	-3.40	-4.77	-11.33	-9.19	-13.31	-16.12	-15.52	-15.82	-19.93	-30.45	-31.97
1923 5	1053.3	-3.45	-1.55	-1.65	-18.64	-13.13	-1.83	-4.68	-8.48	-8.39	-12.47	-13.23	-16.65
1926 10	3567.3*	-.51	-3.51	-.63	-1.47	5.68	3.91	-.54	3.19	-.71	3.02	5.06	.02
1929 8	3911.0	-2.91	-1.61	-4.19	-9.92	-9.77	-7.24	-9.97	-9.41	-4.99	-10.94	-8.62	-17.57
1937 5	2301.7*	-.98	-6.11	-8.33	-9.72	-13.02	-10.06	-11.24	-15.76	-16.02	-14.15	-16.89	-13.93
1945 2	1397.7*	-.83	6.46	15.48	18.70	22.85	41.31	72.36	113.21	129.10	143.62	211.02	229.48
1948 11	7093.0*	.61	-2.23	-5.58	-4.72	-6.78	-1.90	-1.56	-5.65	4.64	8.78	2.06	10.74
1953 7	8595.7	-5.32	-2.83	1.24	-1.67	4.76	-2.61	4.49	8.24	8.64	.90	9.04	14.02
1957 7	11786.0	4.78	.35	.48	-8.54	-9.37	-2.65	-6.52					

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	1312.7	-29.76	-25.57	-26.41	-29.00	-28.77	-30.37	-23.74	-26.10	-31.13	-25.04	-30.60	-29.53
1923 5	1053.3	-17.12	-15.51	-20.06	-7.82	-6.01	-16.93	-1.46					
1926 10	3567.3*	7.11	4.25	4.36	11.85	6.21	5.26	16.25	13.67	7.90	9.30	1.76	14.96
1929 8	3911.0	-7.59	-8.31	-10.94	-8.36	-14.14	-12.63	-8.18	-9.31	-11.89	-8.33	-7.80	-12.91
1937 5	2301.7*	-14.11	-13.63	-9.37	-13.54	-17.58	-13.11	-14.45	-14.67	-14.15	-9.15	-17.06	-2.33
1945 2	1397.7*	219.25	229.98	228.26	207.80	206.67	191.91	173.24	162.87	141.19	140.33	136.75	137.25
1948 11	7093.0*	11.88	12.17	14.80	14.53	7.34	21.43	20.12	5.61	10.35	-.56	.65	2.55
1953 7	8595.7	14.32	20.99	20.65	31.69	40.79	34.51	39.22	38.13	24.33	30.79	41.00	32.01
1957 7	11786.0												

LIABILITIES OF BUSINESS FAILURES

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	78.0	96	101	127	116	418	255	299	374	437	305	467	431
1923 5	433.0	363	415	361	363	891	498	410	424	352	775	470	394
1926 10	359.0	324	362	424	460	475	512	406	437	501	413	415	407
1929 8	405.0	437	364	536	544	416	462	551	522	566	686	485	559
1937 5	129.0*	144	142	153	115	152	178	257	181	204	363	192	184
1945 2	3613.3*	3538	950	2260	3370	3620	1220	2100	3110	1310	1450	4330	2840
1948 11	2669.0	3111	2038	2813	3099	3100	2508	3028	2295	3149	2219	2193	2478
1953 7	3516.7	3206	3800	4030	4000	4207	2873	4083	5114	4168	3666	4161	3323
1957 7	4867.0	4889	5103	5155	5750	4358	6256	5581					

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	78.0	597	535	371	614	438	498	452	468	597	530	694	610
1923 5	433.0	432	428	581	434	406	308	360	450	393	270	358	398
1926 10	359.0	357	406	393	442	435	365	388	377	344	613	430	393
1929 8	405.0	601	655	570	675	644	537	586	541	545	562	744	602
1937 5	129.0*	179	164	167	168	142	134	338	189	122	163	169	150
1945 2	3613.3*	4020	3670	3730	3170	3400	3960	6180	6400	9800	15000	1504	1236
1948 11	2669.0	1887	2813	2261	2325	2063	2313	1943	2057	1864	1640	1487	2050
1953 7	3516.7	3661	4088	3152	3812	3856	3677	3595	3679	3526	3306	3667	3355
1957 7	4867.0												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	78.0	23.08	29.49	62.82	48.72	435.90	226.92	283.33	379.49	460.26	291.03	498.72	452.56
1923 5	433.0	-16.17	-4.16	-16.63	-16.17	105.77	15.01	-5.31	-2.08	-18.71	78.98	8.54	-9.01
1926 10	359.0	-9.75	.84	18.11	28.13	32.31	42.62	13.09	21.73	39.55	15.04	15.60	13.37
1929 8	405.0	7.90	-10.12	32.35	34.32	2.72	14.07	36.05	28.89	39.75	69.38	19.75	38.02
1937 5	129.0*	11.63	10.08	18.60	-10.85	17.83	37.98	99.22	40.31	58.14	181.40	48.84	42.64
1945 2	3613.3*	-2.31	-73.71	-37.45	-6.73	.18	-66.24	-41.88	-13.93	-63.75	-59.87	19.83	-21.40
1948 11	2669.0	16.56	-23.64	5.39	16.11	16.15	-6.03	13.45	-14.01	17.98	-17.01	-20.08	-7.16
1953 7	3516.7	-8.83	8.06	14.60	13.74	19.63	-18.30	16.10	45.42	18.52	4.25	18.32	-5.51
1957 7	4867.0	.45	4.85	5.92	18.14	-10.46	28.54	14.67	.00	.00	.00	.00	.00

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	78.0	665.38	585.90	375.64	687.18	461.54	538.46	479.49	500.00	665.38	579.49	789.74	682.05
1923 5	433.0	-.23	-1.15	34.18	.23	-6.24	-28.87	-16.86	3.93	-9.24	-37.64	-17.32	-8.08
1926 10	359.0	-.56	13.09	9.47	23.12	21.17	1.67	8.08	5.01	-4.18	70.75	19.78	9.47
1929 8	405.0	48.39	61.73	40.74	66.67	59.01	32.59	44.69	33.58	34.57	38.77	83.70	48.64
1937 5	129.0*	38.76	27.13	29.46	30.23	10.08	3.88	162.01	46.51	-5.43	26.36	31.01	16.28
1945 2	3613.3*	11.23	1.57	3.23	-12.27	-5.90	9.59	71.03	77.12	171.22	315.13	-58.38	-65.79
1948 11	2669.0	-29.38	5.39	-15.29	-12.89	-22.71	-13.34	-27.20	-22.93	-30.16	-38.55	-44.29	-23.19
1953 7	3516.7	4.18	16.25	-10.37	8.40	9.65	4.56	2.23	4.62	.26	-5.99	4.27	-4.60
1957 7	4867.0												

BASIC PRICES

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	20572.7	20832	20754	20723	20304	19614	19090	18601	17442	16292	14650	13144	12516
1923 5	13520.0	13237	12955	12867	13006	13118	13287	13354	13234	13046	12776	12607	12425
1926 10	12754.7	12760	12802	12668	12536	12542	12486	12433	12403	12482	12742	13084	13330
1929 8	12630.0	12684	12552	12319	11958	11597	11369	11204	11061	10856	10668	10506	10458
1937 5	1471.7*	1414	1444	1419	1382	1254	1139	1105	1120	1093	1074	1027	995
1945 2	1833.3	1836	1837	1838	1837	1836	1839	1844	1857	1868	1871	1873	1880
1948 11	1037.0*	1014	965	901	856	801	802	774	785	840	848	804	808
1953 7	878.3	886	897	861	877	879	883	879	896	929	931	916	915
1957 7	898.0	894	875	846	848	846	846	857	857				

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	20572.7	12117	11620	11098	10719	10672	10893	11072	11138	11270	11332	11342	11395
1923 5	13520.0	12260	12424	12716	12904	13174	13440	13732	13910	13860	13760	13501	13468
1926 10	12754.7	13464	13554	13550	13434	13378	13427	13314	13166	13166	13236	13204	13070
1929 8	12630.0	10362	10178	9946	9672	9408	9239	9199	9071	8778	8709	8784	8642
1937 5	1471.7*	984	1056	1038	1036	1051	1049	1036	1039	1035	1041	1018	1034
1945 2	1833.3	1897	1903	1918	1972	2402	2414	2405	2529	2904	3010	3039	3156
1948 11	1037.0*	805	827	821	812	826	875	890	989	1058	1147	1144	1219
1953 7	878.3	912	906	902	911	891	910	915	889	902	893	906	913
1957 7	898.0												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	20572.7	1.26	.88	.73	-1.31	-4.66	-7.21	-10.56	-15.22	-20.81	-28.79	-36.11	-39.16
1923 5	13520.0	-2.09	-4.18	-4.83	-3.80	-2.97	-1.72	-1.23	-2.11	-3.51	-5.50	-6.75	-8.10
1926 10	12754.7	.04	.37	-.68	-1.71	-1.67	-2.11	-2.52	-2.76	-2.14	-.10	2.58	4.51
1929 8	12630.0	.49	-.62	-2.46	-5.32	-8.18	-9.98	-11.29	-12.42	-14.05	-15.53	-16.82	-17.36
1937 5	1471.7*	-3.92	-1.88	-3.85	-6.09	-14.79	-22.60	-24.91	-23.90	-25.73	-27.02	-30.21	-32.39
1945 2	1833.3	.15	.20	.25	.20	.15	.31	.58	1.29	1.89	2.05	2.16	2.55
1948 11	1037.0*	-2.22	-6.94	-13.11	-17.45	-22.76	-22.66	-25.36	-24.30	-19.00	-18.23	-22.47	-22.08
1953 7	878.3	.87	2.13	-1.97	-.15	.08	.53	.08	2.01	5.77	6.00	4.29	4.17
1957 7	898.0	-.44	-2.56	-5.79	-5.90	-5.79	-5.79	-4.57	-4.57				

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	20572.7	-41.10	-43.52	-46.05	-47.90	-48.12	-47.05	-46.18	-45.86	-45.22	-44.92	-44.87	-44.61
1923 5	13520.0	-9.32	-8.11	-5.95	-4.56	-2.56	-.59	1.57	2.88	2.51	1.77	-.14	-.38
1926 10	12754.7	5.56	6.27	6.24	5.33	4.89	5.27	4.38	3.22	3.22	3.77	3.52	2.47
1929 8	12630.0	-17.96	-19.41	-21.25	-23.42	-25.51	-26.85	-27.17	-28.18	-30.50	-31.04	-30.45	-31.58
1937 5	1471.7*	-33.14	-28.24	-29.47	-29.60	-28.58	-28.72	-29.60	-29.40	-29.67	-29.26	-30.83	-29.74
1945 2	1833.3	3.47	3.80	4.62	7.56	31.02	31.67	31.18	37.62	58.40	64.18	65.76	72.15
1948 11	1037.0*	-22.37	-20.25	-20.83	-21.70	-20.35	-15.62	-14.18	-4.63	2.02	10.61	10.32	17.55
1953 7	878.3	3.83	3.15	2.69	3.72	1.44	3.60	4.17	1.21	2.69	1.67	3.15	3.72
1957 7	898.0												

INDEX OF INDUSTRIAL STOCK PRICES

STANDINGS — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	10273.3	9650	9795	9945	9110	9170	9070	8525	8650	8485	7930	7220	7475
1923 5	9620.0	9255	8930	9045	9075	8815	9065	9410	9780	9880	9560	9195	9040
1926 10	15595.0	15395	15930	15465	15815	16010	16405	16880	16885	17535	18385	19330	18980
1929 8	35431.7	36235	29150	22820	24720	25565	26740	27825	28350	26670	24315	22980	22880
1937 5	17328.3	17030	17785	18295	16025	13990	12475	12435	12725	12545	11475	11100	11355
1945 2	15586.7	15690	16022	16608	16674	16400	16792	17780	18506	18934	19244	19812	19650
1948 11	18036.0	17557	17828	17574	17606	17520	17250	16488	17227	17943	18066	18617	19080
1953 7	26980.3	27221	26190	27073	27710	28119	28664	29213	29915	31092	32286	32791	34127
1957 7	50262.7	48797	47179	44338	43673	43694	44568	44416	45014				

STANDINGS — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	10273.3	7570	7505	7700	7670	6920	6860	6695	6935	7170	7570	7980	8060
1923 5	9620.0	9330	9925	10355	10290	10165	10765	11545	12155	12045	12035	11970	12355
1926 10	15595.0	18995	19800	19895	19535	20450	21245	21630	21150	21085	22725	23930	24745
1929 8	35431.7	22500	19875	18095	17215	16725	18155	18005	15800	14145	13930	14535	13980
1937 5	17328.3	12280	13975	14095	13650	14865	15225	15110	14565	14485	14205	12710	13300
1945 2	15586.7	19451	20375	20658	20600	20139	19686	17318	16953	16898	17291	17620	18086
1948 11	18036.0	19662	19979	20346	20630	21267	21936	22102	20530	21660	22321	22932	22938
1953 7	26980.3	34606	35271	35829	37575	39384	39843	41025	40891	42299	42155	44079	46216
1957 7	50262.7												

PER CENT CHANGES — FIRST YEAR AFTER PEAK

DATE	AT PEAK	+ 1 MO	+ 2 MO	+ 3 MO	+ 4 MO	+ 5 MO	+ 6 MO	+ 7 MO	+ 8 MO	+ 9 MO	+10 MO	+11 MO	+12 MO
1920 1	10273.3	-6.07	-4.66	-3.20	-11.32	-10.74	-11.71	-17.02	-15.80	-17.41	-22.81	-29.72	-27.24
1923 5	9620.0	-3.79	-7.17	-5.98	-5.67	-8.37	-5.77	-2.18	1.66	2.70	-.62	-4.42	-6.03
1926 10	15595.0	-1.28	2.15	-.83	1.41	2.66	5.19	8.24	8.27	12.44	17.89	25.23	21.71
1929 8	35431.7	2.27	-17.73	-35.59	-30.23	-27.85	-24.53	-21.47	-19.42	-24.73	-31.37	-35.14	-35.42
1937 5	17328.3	-1.72	2.64	5.58	-7.52	-19.26	-28.01	-28.24	-26.56	-27.60	-33.78	-35.96	-34.47
1945 2	15586.7	.66	2.79	6.55	6.98	5.22	7.73	14.07	18.73	21.48	23.46	27.11	26.07
1948 11	18036.0	-2.66	-1.15	-2.56	-2.38	-2.86	-4.36	-8.58	-4.49	-.52	.17	3.22	5.79
1953 7	26980.3	.89	-2.93	.34	2.70	4.21	6.24	8.27	10.88	15.24	19.66	21.54	26.49
1957 7	50262.7	-2.92	-6.13	-11.79	-13.11	-13.07	-11.33	-11.63	-10.44				

PER CENT CHANGES — SECOND YEAR AFTER PEAK

DATE	AT PEAK	+13 MO	+14 MO	+15 MO	+16 MO	+17 MO	+18 MO	+19 MO	+20 MO	+21 MO	+22 MO	+23 MO	+24 MO
1920 1	10273.3	-26.31	-26.95	-25.05	-25.34	-32.64	-33.22	-34.83	-32.50	-30.21	-26.31	-22.32	-21.54
1923 5	9620.0	-3.01	3.17	7.64	6.96	5.67	11.90	20.01	26.35	25.21	25.10	24.43	30.51
1926 10	15595.0	21.80	26.96	27.57	25.26	31.13	36.23	38.70	35.62	35.20	45.72	53.45	58.67
1929 8	35431.7	-36.50	-43.91	-48.93	-51.41	-52.80	-48.76	-49.18	-55.41	-60.08	-60.68	-58.98	-60.54
1937 5	17328.3	-29.13	-19.35	-18.66	-21.23	-14.22	-12.14	-12.80	-15.93	-16.41	-18.02	-26.65	-23.25
1945 2	15586.7	24.79	30.72	32.54	32.16	29.21	26.30	11.11	8.77	8.41	10.93	13.04	16.03
1948 11	18036.0	9.02	10.77	12.81	14.38	17.91	21.62	22.54	13.83	20.09	23.76	27.15	27.18
1953 7	26980.3	28.26	30.73	32.80	39.27	45.97	47.67	52.05	51.56	56.78	56.24	63.37	71.29
1957 7	50262.7												

SOURCE NOTES AND UNITS

Series and Period Covered		Unit	Source
1. Nonagricultural Employment	1929–39	10,000 persons	BLS and FRB
	1945–58	1,000 persons	BLS
2. Unemployment Rate	1929–39	Per 10,000 in labor force	NICB and NBER
	1947–58	Per 10,000 in labor force	Census
3. Gross National Product	1923–29	$10 million, annual rate	Barger and Klein
	1945–58	Million dollars, annual rate	OBE
4. Industrial Production	1920–58	1947–49 = 100	FRB
5. Freight Cars Loaded	1920–58	1,000 cars per week	AAR
6. Bank Debits outside N.Y.C.	1920–39	$100 million	FRB, 140 centers outside N.Y.C.
	1945–58	$100 million	FRB, 344 centers outside N.Y.C.
7. Personal Income	1923–28	$10 million	Barger and Klein
	1929–58	$100 million, annual rate	OBE
8. Retail Store Sales	1920–31	1947–49 = 100	FRB (department store sales only)
	1937–50	Billion dollars	OBE
	1953–58	Billion dollars	Census
9. Wholesale Prices, ex Farm Products and Food	1920–25	1926 = 1,000	BLS
	1926–58	1947–49 = 1,000	BLS
10. Corporate Profits after Taxes	1920–39	$10 million, annual rate	Barger
	1945–58	$100 million, annual rate	OBE
11. Average Workweek in Manufacturing	1923–31	1/10 hour per week	NICB, 25 mfg. industries
	1937–58	1/10 hour per week	BLS, all mfg. industries
12. Layoff Rate, Manufacturing	1920–31	Per 10,000 employed	Metrop. Life Ins. Co. and BLS
	1937–58	Per 10,000 employed	BLS
13. Gross Accession Rate, Manufacturing	1920–31	Per 10,000 employed	Metrop. Life Ins. Co. and BLS
	1937–58	Per 10,000 employed	BLS
14. New Orders for Durable Goods	1920–28	1923–25 = 1,000	NBER
	1929–39	1935–39 = 100	NICB
	1945–58	Million dollars	OBE
15. Residential Construction Contracts	1920–22	10,000 square feet	Dodge, 27 eastern states
	1923–25	10,000 square feet	Dodge, 36 eastern states
	1926–55	10,000 square feet	Dodge, 37 eastern states
	1957–58	10,000 square feet	Dodge, 48 states
16. Commercial and Industrial Construction Contracts	1920–25	10,000 square feet	Dodge, 27 eastern states
	1926–28	10,000 square feet	Dodge, 36 eastern states
	1929–55	10,000 square feet	Dodge, 37 eastern states
	1957–58	10,000 square feet	Dodge, 48 states
17. Number of Incorporations	1920–25	Number	Evans, 6 states
	1926–31	Number	Evans, 8 states
	1937–39	Number	Evans, 6 states
	1945–47	Number	Evans, 4 states
	1948–58	Number	Dun and Bradstreet, 48 states
18. Liabilities of Business Failures	1920–31	$10,000	Dun and Bradstreet
	1937–39	$10,000	Dun and Bradstreet
	1945–58	$10,000	Dun and Bradstreet
19. Basic Commodity Prices	1920–31	1/10 cent	Bradstreet, 96 commodities
	1937–47	Aug. 1939 = 1,000	BLS, spot prices, 28 commodities
	1948–50	1947–49 = 1,000	BLS, spot prices, 22 commodities
	1953–58	1947–49 = 1,000	BLS, spot prices, 22 commodities
20. Index of Industrial Stock Prices	1920–28	Cents per share	Dow Jones, 20 indus. common stocks
	1929–58	Cents per share	Barron's, 30 indus. common stocks

BLS: U. S. Bureau of Labor Statistics.
FRB: Board of Governors of the Federal Reserve System.
NICB: National Industrial Conference Board
NBER: National Bureau of Economic Research.
Census: U. S. Bureau of the Census.
Barger and Klein: Harold Barger and Lawrence Klein, unpublished estimates.
OBE: U. S. Department of Commerce, Office of Business Economics.
AAR: Association of American Railroads, Car Service.
Barger: Harold Barger, *Outlay and Income in the United States, 1921–1938* (New York: NBER, 1942).
Metrop. Life Ins. Co.: Metropolitan Life Insurance Company.
Dodge: F. W. Dodge Corporation, Dodge Statistical Research Service.
Evans: George Heberton Evans, Jr., *Business Incorporations in the United States: 1800–1943* (New York: NBER, 1948), Appendix 6.
Dun and Bradstreet: Dun and Bradstreet, Inc. Prior to February 1933: R. G. Dun and Co.
Bradstreet: The Bradstreet Co.
Dow Jones: *The Dow Jones Averages*, 13th Edition (New York: Barron's Publishing Company, Inc., 1948), and "Barron's National Business and Financial Weekly."

APPENDIX B

SOURCES OF CURRENT DATA AND SEASONAL ADJUSTMENT FACTORS

I. SOURCES OF CURRENT DATA

1. *Nonagricultural Employment*
   U. S. Department of Labor, Bureau of Labor Statistics, *Employment and Earnings*, monthly report and advance news release, "Employment, Hours, and Earnings." Seasonal adjustment by compiler.
2. *Unemployment Rate*
   U. S. Department of Commerce, Bureau of the Census, "Monthly Report on the Labor Force" (*Current Population Reports: Labor Force, Series P-57*). Number of unemployed as per cent of civilian labor force. Seasonal adjustment by compiler.
3. *Gross National Product (quarterly)*
   U. S. Department of Commerce, Office of Business Economics, *Survey of Current Business* and advance news release, "National Income and Product." Preliminary estimates are published in *Economic Indicators* (Council of Economic Advisers, prepared for Joint Committee on the Economic Report). Seasonal adjustment by compiler.
4. *Index of Industrial Production*
   Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin* and advance release, "Business Indexes." Seasonal adjustment by compiler.
5. *Freight Carloadings*
   Association of American Railroads, Car Service Division, *Revenue Freight Loaded and Received from Connections*, weekly report. Monthly averages derived by NBER from weekly figures; a week ending on the 1st, 2nd, or 3rd day of the month is included with the weeks for the preceding month. A preliminary estimate of the current monthly average may be obtained by averaging the last four weekly figures. Seasonal adjustment by NBER.
6. *Bank Debits outside New York City*
   Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin* and advance release, "Bank Debits to Demand Deposit Accounts." Derived by NBER by deducting New York City from total for all reporting centers. Seasonal adjustment by NBER.
7. *Personal Income*
   U. S. Department of Commerce, Office of Business Economics, *Survey of Current Business* and advance news release, "Personal Income." Seasonal adjustment by compiler.
8. *Sales of Retail Stores*
   U. S. Department of Commerce, Bureau of the Census. Seasonally adjusted by Office of Business Economics in *Survey of Current Business* and in Bureau of the Census news release, "Advance Retail Sales Report."
9. *Index of Wholesale Prices, All Commodities other than Farm Products and Foods*
   U. S. Department of Labor, Bureau of Labor Statistics, *Wholesale (Primary Market) Price Index*, monthly report. As a preliminary estimate for the current month, the weekly index for the week closest to the middle of the month may be used. (The BLS compiles a weekly index based on about 10 per cent of the commodities quoted monthly, available in weekly news release.) No seasonal adjustment necessary.
10. *Corporate Profits after Tax (quarterly)*
    Same as series 3.
11. *Average Workweek, Manufacturing*
    Same as series 1. Seasonal adjustment by NBER.
12. *Layoff Rate, Manufacturing*
    Same as series 1 and advance news release, "Factory Labor Turnover." Seasonal adjustment by NBER.

II. SEASONAL ADJUSTMENT FACTORS FOR SERIES ADJUSTED BY NBER

Series	Year	Month											
		Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sep.	Oct.	Nov.	Dec.
Freight carloadings	1957–58	92	92	93	98	102	105	100	107	108	112	101	89
Bank debits outside N.Y.C.[a]	1957	107	91	103	100	101	97	100	95	98	106	97	106
	1958	107	91	103	100	98	101	100	93	98	106	94	109
Average workweek, mfg.	1957–58	99.8	99.9	99.9	99.2	99.5	99.9	99.4	100.0	100.2	100.4	100.5	101.3
Layoff rate, mfg.	1957–58	110	99	97	100	99	89	97	100	92	96	106	115
Gross accession rate, mfg.	1957–58	95	89	95	94	100	124	100	114	116	112	91	70
Residential building contracts, 37 states[b]	1957	73.1	85.4	105.3	120.7	121.9	108.3	116.0	111.3	100.3	101.3	86.5	70.0
	1958	72.5	86.0	104.7	120.2	122.3	108.6	116.6	113.2	101.1	101.1	86.4	68.1
Residential building contracts, 11 states[b]	1957–58	82	87	117	116	109	111	104	109	93	101	89	81
Commercial building contracts, 48 states[c]	1957–58	90	79	98	107	117	108	110	97	93	108	91	101
Industrial building contracts, 48 states[c]	1957–58	77	81	85	115	106	110	118	105	98	113	86	107
New incorporations, number	1957–58	114	95	113	110	107	104	96	92	89	95	86	99
Liabilities of business failures	1957–58	103	117	112	102	105	100	97	89	89	92	92	104

Note: These seasonal factors are currently being used at the National Bureau. However, in applying them to future data they should be continuously reviewed and revised when found to be no longer applicable.

[a] Index computed to take account of leap years and the varying number of Saturdays and Sundays in each month.

[b] The seasonally adjusted data for 37 states and for 11 states are totaled to obtain seasonally adjusted data for 48 states. The 11 state factors are based on BLS permit valuations.

[c] The seasonally adjusted data for commercial building contracts and for industrial building contracts are totaled to obtain seasonally adjusted data for commercial and industrial building contracts.

13. *Gross Accession Rate, Manufacturing*
Same as series 12. Seasonal adjustment by NBER.

14. *Manufacturers' New Orders, Value, Durable Goods Industries*
U. S. Department of Commerce, Office of Business Economics, *Survey of Current Business* and advance news release, "Manufacturers' Sales and Inventories." Seasonal adjustment by compiler.

15. *Construction Contracts Awarded, Residential Building, Floor Area*
F. W. Dodge Corporation, Dodge Statistical Research Service, *Construction Contracts: United States Summary,* monthly report. Seasonal adjustment by NBER.

16. *Construction Contracts Awarded, Commercial and Industrial Buildings, Floor Area*
Same as series 15. Data are sum of "Commercial Buildings" and "Manufacturing Buildings." Seasonal adjustment by NBER.

17. *New Business Incorporations, Number*
Dun & Bradstreet, Inc., *Dun's Statistical Review* through 1957. Press release, "New Business Incorporations" thereafter. Seasonal adjustment by NBER.

18. *Current Liabilities of Industrial and Commercial Failures*
Dun & Bradstreet, Inc., *Dun's Statistical Review* through 1957. Press release, "Monthly Business Failures" thereafter. Seasonal adjustment by NBER.

19. *Index of Spot Market Prices*
U. S. Department of Labor, Bureau of Labor Statistics, "Daily Index Numbers and Spot Primary Market Commodity Prices," weekly release; also daily in *New York Times.* Data used are for the trading day on or nearest the 15th of each month. No seasonal adjustment necessary.

20. *Index of Industrial Stock Prices, Dow-Jones*
Barron's National Business and Financial Weekly; also daily in *Wall Street Journal.* Monthly average of daily closing prices of thirty common stocks. As a preliminary estimate of the current monthly average, the average for the preceding 30 or 31 days may be used. No seasonal adjustment necessary.

THE ACCURACY AND STRUCTURE OF INDUSTRY EXPECTATIONS IN RELATION TO THOSE OF INDIVIDUAL FIRMS*

ROBERT FERBER
University of Illinois

In this empirical study of bimonthly labor force expectations of Chicago area employers, considerable disparities were obtained between the accuracy and structure of the expectations at the level of the firm and at the level of the corresponding industry aggregates. The results lend support to the hypothesis that the structure of individual firm anticipations may be wholly dissimilar to that of the aggregate. Considerable reliance on mechanical projection of level is apparent among the individual firms. At the same time, there is evidence that firms which make some effort to anticipate employment are more accurate than other firms, judging by naive model tests.

THE extent to which results obtained for aggregates are also valid for the individual units that make up the aggregates is one of the central questions of empirical analysis. The question is of particular relevance to business behavior because the available data are generally in the form of aggregates and it is desired to gauge the applicability of the results to individual firms. The area of businessmen's expectations is a case in point: firms are reluctant to make public their own expectations so that such data are likely to be released only in the form of aggregates. To impute characteristics of individual firm expectations is in many ways a hopeless task, considering the numerous more or less arbitrary assumptions involved in the disaggregation process.

The present study represents an empirical exploration of the extent to which findings pertaining to the accuracy and structure of individual firm expectations compare with those that would be obtained for the corresponding industry aggregates. The study is made possible by the availability of monthly labor force expectations, and the corresponding actual figures, for Chicago area employers on a fairly continuous basis for several postwar years. By carrying out the analysis first on the individual firms and then on the aggregates resulting when the firms are combined into industry groups, comparisons can be made between the two sets of findings.

THE DATA

Since 1946 Illinois employers have been asked by the Illinois Department of Labor (as part of a larger program supervised by the U. S. Bureau of Employment Security) to report monthly the anticipated size of their labor force two months in advance and four months in advance as well as the actual size of the

* This study is an outgrowth of the project on "Expectations and Business Fluctuations" financed by the Merrill Foundation and conducted at the University of Illinois under the direction of Franco Modigliani. The author would like to thank the University Research Board and the Bureau of Economic and Business Research of the University of Illinois for supplemental assistance. He is also indebted to Franco Modigliani for helpful suggestions, to Nai Ruenn Chen and Fadil Zuwaylif of the Bureau staff for aid with the statistical analysis, and, last, but not least, to Elizabeth Slotkin and Stanley Neyhaus of the Illinois Department of Labor for their assistance in securing the basic data.

labor force in the given month. The only other data requested in this report relate to separations from, and additions to, the payroll during the given month. The reports are sent out about the middle of the month, are to be returned by the end of the month, and are published bimonthly by industry groupings.

Coverage. Cooperation is sought of all business firms in the Chicago-Calumet Labor Market Area with 65 or more employees. In terms of sales volume, employment or other indicators of economic importance, a substantial rate of return is achieved because special efforts are made to secure returns from the larger firms in each industry.

Representativeness of replies. The extent to which the anticipations represent the official outlook of the firm, if such exists, and the level of operations from which these anticipations emanate is difficult to estimate, since no questioning of firms could be conducted. To judge by the title of the person to whom data requests are sent, it is the personnel manager or his assistant who supplies the data in most instances, though occasionally a production manager or a vice-president enter the picture. However, quite a few returns are sent in with the signature of someone other than the addressee, usually a person of lower rank, often a clerk. This does not necessarily mean that the clerk works up the data. In the opinion of Department of Labor analysts, the personnel manager or another higher official is invariably consulted in such cases.

Meaning of anticipations. The anticipations reported by the respondents are not strictly forecasts in the sense of predicting what is likely to happen. Rather they are more in the nature of conditional expectations. Thus, the wording of the question has been as follows:

Provided an adequate supply of labor will be available, what will your total employment be on:

a. (15th of month two months hence)?_____

b. (15th of month four months hence)? _____

In periods of labor shortages, as in the immediate postwar years or during and shortly after a strike, these anticipations are, in a sense, conditional intentions; they are clearly not predictions of what is expected to happen and are of little value insofar as gauging actual business conditions is concerned. On the other hand, as long as manpower supplies are adequate, the anticipations would seem to be equivalent to predictions of future employment. This is true for nearly all of the period covered by the present study, which is Fall 1951 to the end of 1955, the principal exceptions being isolated cases of labor stoppages.

Data used. The bulk of the analysis is carried out for 59 principal firms (plants)[1] in 14 industries in the Chicago-Calumet Labor Market Area. These are predominantly large firms because relatively complete records are available only for such firms, and it is these firms which have the greatest impact on economic activity. Restriction of the study to the Chicago area was necessitated by the lack of data for other areas not to mention other considerations as well, e.g., efficiency of operation. At the same time, resources did not permit

[1] The two terms are used interchangeably in this study though more than one plant of some firms was included in the sample. Insofar as hiring and employment expectations are concerned, these plants clearly operated independently of each other.

intensive study of more than a few industries, and these were selected with a view toward (a) availability of almost continuous reports for individual companies for almost four years, and (b) representation in the study of industries differing with respect to such characteristics as seasonal variations, susceptibility to cyclical influences, and growth trends. On the basis of these considerations, the following industries were selected for study from those eligible. The number of firms studied in each industry is shown in parentheses; the period of observation is shown at the right.

Food, excl. meat packing (4)	May 1952–Jan. 1956
Meat packing (3)	March 1952–Jan. 1956
Printing and publishing (3)	March 1952–Jan. 1956
Iron and steel (12)	Nov. 1951–Jan. 1956
Cans (2)	July 1952–Jan. 1956
Other fabricated metals (3)	May 1952–Jan. 1956
Nonelectrical machinery (6)	May 1952–Jan. 1956
Farm machinery (3)	Nov. 1951–Jan. 1956
Electrical machinery (3)	Sept. 1951–Jan. 1956
Communication equipment (6)	May 1952–Jan. 1956
Automobiles and parts (3)	Nov. 1952–Jan. 1956
Other transportation equipment (5)	March 1952–Jan. 1956
Utilities (2)	March 1952–Jan. 1956
Retail trade (4)	March 1952–Jan. 1956

COMPARATIVE ACCURACY

There are two principal vantage points from which the accuracy of a set of expectations can be considered. One is their accuracy as forecasts of level, namely, measures of the accuracy of the anticipations themselves without regard to any benchmarks of "good" or "bad" estimates. Distinct from this type of evaluation is the accuracy of the anticipations as forecasts of change, an approach which eliminates to a large extent the inherent serial correlation in economic data.[2]

Two aspects of the anticipations as forecasts of change are considered: their superiority relative to forecasts of no change, and their usefulness in providing information on forthcoming change. In the first sense, the accuracy of the anticipations is compared with the approach of naive models, set up to predict the actual movement on the basis of some simple mechanical extrapolation formula. If the accuracy of the anticipations is no better than simple mechanical projection, the anticipations are not likely to be of much value for direct forecasting purposes, though they may still possess value in other respects. There remains the possibility, however, that the anticipations may nevertheless throw light on impending change, a possibility that is explored by means of a correlation approach.

Forecasts of Level

The average absolute per cent error of the two-month and four-month anticipations is presented in Table 320 on two bases of classification. On the one base are presented the averages of the errors of the individual firms in the industry

[2] Admittedly, these are not the only criteria of adequacy of a set of anticipations. Another facet, which could not be explored in this study, is the extent to which revisions in anticipations, and employment, are sensibly related to disappointments in plans.

TABLE 320

AVERAGE ABSOLUTE PER CENT ERROR OF EMPLOYERS' ANTICIPATIONS

Measure of error	Two-month anticipations		Four-month anticipations	
	Average of firms' errors	Average of industry errors	Average of firms' errors	Average of industry errors
Average absolute per cent error	4.7%	3.2%	7.0%	5.0%
Standard deviation of errors	7.0	3.5	9.9	5.7
Coefficient of variation	148.9	109.4	141.4	114.0
Average absolute per cent error by trend of employment				
Rising trend	5.7	3.3	7.4	4.9
Level trend	2.4	2.1	3.3	3.5
Falling trend	8.8	6.5	14.2	9.7

as against the average absolute error that is obtained *after* the individual firm anticipations have been aggregated. The latter is the error that would be obtained if expectations data were reported only for the industry as a whole, as is true for the railroad shippers' forecasts and most other types of data. For the former series, the errors of the individual firms in an industry are averaged.

The second classification is based on the direction of movement of actual employment in the firm or industry, as the case may be, at the time to which the anticipation refers. Three trends are distinguished in the movement of employment: rising, level and falling. The definition of these trends depends on the extent of seasonal variation present in actual employment for each industry. If the relative amplitude of the seasonal index computed for the particular industry is small, by which we shall mean less than 10 per cent, the trends are defined with reference to A_t/A_{t-2}. The figure for a particular month is said to exhibit a falling trend if A_t/A_{t-2} is less than .97, a level trend if A_t/A_{t-2} is between .97 and 1.03 inclusive, and a rising trend if A_t/A_{t-2} exceeds 1.03.

On the other hand, if the seasonal index for a particular industry contains substantial fluctuations—meaning that the relative amplitude of the series is 10 per cent or more—the trend in employment is defined with reference to the change between the current month and the corresponding month of the preceding year, i.e., A_t/A_{t-12}. This is true of six industries and of 22 sample firms.[3] In these cases, a 5 per cent change was selected as the boundary point for differentiating between trends, the trend being a rising one if A_t/A_{t-12} exceeded 1.05, falling if A_t/A_{t-12} was less than .95, and level otherwise.

The rationale back of this procedure is that consideration is likely to be given to seasonal fluctuations in anticipating future trends only where substantial

[3] Two firms exhibited no seasonal pattern in employment in contrast to sharp seasonal patterns by the other firms in these industries (communications equipment, and printing and publishing), and one firm exhibited a sharp seasonal while the other firms in that industry (nonelectrical machinery) did not. These firms were treated separately from the others in each industry, with no seasonal adjustment for the first two and computation of a separate seasonal pattern for the third firm.

seasonal fluctuations are present. The distinction between "moderate" fluctuations and "substantial" ones is admittedly, and necessarily, an arbitrary one—as are the breaking points between the three types of trends in employment—but it does nevertheless serve the purpose at hand.

A number of significant points are evident in Table 320 with regard to the accuracy of these anticipations data. Most relevant to the present discussion, however, is the fact that the average error of the anticipations on the industry-wide basis is considerably less than the corresponding average for the individual firms, the differential being of the order, roughly, of 40 to 50 per cent. The same tendency is apparent for the variability in the errors, including the relative variability, and when the errors are segregated by trend in the labor force.

This phenomenon may reflect either or both of two possibilities. One is that the process of combining individual firm anticipations permits for greater accuracy through cancellation of opposing errors. The other possibility is that it may reflect the greater accuracy of the larger firms' anticipations. Exploring these alternatives reveals that it is mainly the cancellation of opposing errors rather than superior accuracy of the larger firms' anticipations that accounts for the lower error of the anticipations on an industry-wide basis.[4]

Similar results are obtained when the same comparisons are made for each of the 14 industry classifications used in this study.[5] Though the accuracy of the firms' anticipations varies substantially by industry, the accuracy of the industry-wide anticipations is invariably more than that of the corresponding firm anticipations, particularly so in the heavy goods industries such as iron and steel, the machinery industries, and other transportation equipment. The differential is low or non-existent, on the other hand, in meat packing and utilities.

It is noteworthy that industries with high average errors for individual firm anticipations also have high average errors of anticipations on an industry-wide basis. The rank correlation coefficient between the two types of errors is .91 for the two-month anticipations and .70 for the four-month anticipations.

Forecasts of Change

Naive Model Tests. The ability of the anticipations to forecast change can be evaluated, from a practical point of view, only when we have some standard against which to assess their accuracy. Essentially, what we ask is how "good" or how "bad" the anticipations may be relative to other means of predicting employment. In practice, this is likely to be the most important consideration, for if the anticipations provide the best available means of forecasting employment they may be a highly useful set of data irrespective of the closeness with which they approximate actual employment. On the other hand, if the anticipations are less accurate than some mechanical, or "naive," forecasting approach, they may not possess much practical value even if they prove to be very close to actual employment.

Two "naive models" are used to test the external accuracy of the anticipations. One is the simple projection of the level of the current labor force figure;

[4] The full results are contained in a forthcoming monograph to be published by the Bureau of Economic and Business Research of the University of Illinois.

[5] *Ibid.*

the other is the projection of the current figure modified by the most recent rate of change in employment. The actual formulas used depend on the extent to which seasonal variation is present in the employment data. If relatively small seasonal variations are present in the data, the two "naive" projections for month t can be represented as follows for the two-month anticipations:

$$E_{t+2} = A_t$$

and

$$E_{t+2} = A_t \frac{A_t}{A_{t-2}}$$

Similar expressions apply for the four-month anticipations.

Where seasonal variation is present, two procedures can be followed. One is to use the above expressions after adjusting the employment figures for seasonal variations. The other, which is used here, is to make implicit adjustment for seasonal variation by extrapolating the employment figure for the corresponding month of the preceding year, with or without allowance for change in employment during the intervening period.[6] By this procedure, the two naive models for the two-month anticipations becomes:[7]

$$E_{t+2}' = A_t \frac{A_{t-10}}{A_{t-12}}$$

$$E_{t+2}'' = A_t \frac{A_{t-10}}{A_{t-12}} \left[1 + \frac{1}{6} \left\{ \frac{A_t}{A_{t-12}} - 1 \right\} \right]$$

and for the four-month anticipations:

$$E_{t+4}' = A_t \frac{A_{t-8}}{A_{t-12}}$$

$$E_{t+4}'' = A_t \frac{A_{t-8}}{A_{t-12}} \left[1 + \frac{1}{3} \left\{ \frac{A_t}{A_{t-12}} - 1 \right\} \right]$$

The extent to which the labor force anticipations are more accurate than these naive models is shown in Table 323 for firms and for industries. The firms are coded in the table in descending order by size of employment, e.g., A1 is the largest employer in the sample from the food industry and A4 is the smallest. The table indicates that 30 to 40 per cent of the firms (24 or 18 out of 59,

---

[6] A general-purpose model could be set up by using seasonal coefficients, as follows:

$$E_{t+2} = K_1 A_t (S_{t+2}/S_t) + K_2 A_{t-2} (S_t/S_{t-2}),$$

with K_1 and K_2 being predetermined constants.

This approach is not particularly "naive," however, besides raising the problem of estimating seasonal coefficients, and was not used here.

[7] A_{t-10}/A_{t-12} in E'_{t+2} and E''_{t+2} is essentially an adjustment of the change in seasonal pattern from month t to $t+2$, on the assumption of a constant year-to-year seasonal. The same is true of A_{t-8}/A_{t-12} for the four-month anticipations.

TABLE 323

ACCURACY OF LABOR FORCE ANTICIPATIONS RELATIVE TO TWO NAIVE MODELS

Industry	Firm†	Naive model used‡	Proportion of times accuracy of two-month anticipations equals or exceeds		Proportion of times accuracy of four-month anticipations equals or exceeds	
			E_{t+2}'	E_{t+2}''	E_{t+4}'	E_{t+4}''
Food, excl. meat packing	A1	1	43%	80%	19%	53%
	A2	1	62	52	57	74*
	A3	1	36	38	43	79*
	A4	1	100*	76*	100*	74*
	Industry	1	52	65	38	68
Meat packing	B1	1	55	83*	74*	88*
	B2	1	55	59	76*	56
	B3	1	90*	67	85*	84*
	Industry	1	65	88*	63	79*
Printing and publishing	C1	1	65	53	50	63
	C2	2	41	41	44	44
	C3	2	81*	88*	60	60
	Industry	2	56	56	50	50
Iron and steel	*D1*	1	39	48	41	26
	D2	1	70*	48	82*	63
	D3	1	39	38	36	37
	D4	1	65	50	68*	39
	D5	1	52	62	64	79*
	D6	1	43	52	50	53
	D7	1	56	48	64	58
	D8	1	78*	83*	81*	67
	D9	1	60	67*	75*	64
	D10	1	100*	52	100*	75*
	D11	1	86*	55	80*	61
	D12	1	80*	71*	58	50
	Industry	1	52	40	55	61
Cans	E1	2	53	53	57	64
	E2	2	47	47	29	36
	Industry	2	53	53	43	50
Other fabric metals	F1	1	70*	95*	68	71*
	F2	1	68*	71*	52	74*
	F3	1	30	47	53	65
	Industry	1	65	79*	62	63
Nonelectrical machinery	G1	1	45	57	48	74*
	G2	2	91*	82*	60	70
	G3	1	77*	85*	75*	94*
	G4	1	85*	63	89*	59
	G5	1	86*	62	86*	63
	G6	1	71*	65	68	72*
	Industry	1	63	46	50	69

* Significantly above 50% at .05 probability level.

† Italicized firms are those preparing anticipations as simple mechanical projection of the current level of employment.

TABLE 323—(*continued*)

Industry	Firm†	Naive model used‡	Proportion of times accuracy of two-month anticipations equals or exceeds		Proportion of times accuracy of four-month anticipations equals or exceeds	
			E_{t+2}'	E_{t+2}'	E_{t+4}'	E_{t+4}''
Farm machinery	*H1*	2	43	42	50	60
	H2	2	71	64	50	78*
	H3	2	80*	77*	77*	50
	Industry	2	54	67	75*	67
Electrical machinery	I1	1	62	36	68*	57
	I2	1	62	60	56	57
	I3	1	61	64	33	68
	Industry	1	65	55	73*	50
Communications eqt.	J1	2	75*	69	73*	67
	J2	2	75*	63	80*	80*
	J3	2	75*	75*	87*	100*
	J4	1	43	52	32	53
	J5	2	75*	75*	73*	80*
	J6	2	47	73*	71	62
	Industry	2	73*	73*	64	71
Autos and parts	K1	1	26	44	44	56
	K2	1	100*	89*	100*	75*
	K3	1	61	59	41	75*
	Industry	1	32	50	35	50
Other transp. eqt.	L1	2	94*	88*	88*	94*
	L2	2	55	60	78	71
	L3	2	41	41	69	56
	L4	2	47	71	56	69
	L5	2	94*	88*	81*	93*
	Industry	2	80*	90*	100*	100*
Utilities	*M1*	1	90*	74*	95*	76*
	M2	1	43	41	68*	80*
	Industry	1	35	47	44	69
Retail trade	N1	2	29	35	63	69
	N2	2	67	53	71	79*
	N3	2	53	56	63	73
	N4	2	53	47	50	50
	Industry	2	44	56	63	81*

‡ For industries with slight seasonal variation the two-month model is:

$$E_{t+2}' = A_t; \qquad E_{t+2}'' = A_t \frac{A_t}{A_{t-2}}$$

The four-month model is the same except that A_{t-2} is replaced by A_{t-4}.

2: For industries with appreciable seasonal variation, the two month model is:

$$A_{t+2}' = A_t \frac{A_{t-10}}{A_{t-12}} \qquad E_{t+2}'' = A_t \frac{A_{t-10}}{A_{t-12}} \left[1 + \frac{1}{6} \left\{ \frac{A_t}{A_{t-12}} - 1 \right\} \right]$$

In the four-month model, A_{t-3} replaces A_{t-10} and $\frac{1}{4}$ replaces $\frac{1}{6}$.

depending on the naive model used) were able to anticipate employment significantly better than the naive models two months ahead. Similar results are obtained for the four-month anticipations. On an industry basis, the anticipations did not do as well, proving significantly superior to the naive models for only two or four industries, depending on the naive model used.

The overall average of the proportion of times the anticipations of all firms are more accurate than each naive model, and the overall average of the proportion of times the anticipations for all 14 industries are superior to each naive model, are both in the neighborhood of 60 per cent.

The extent to which relative accuracy is an industry phenomenon is not too clear from these data, for the firms whose anticipations proved generally more accurate than the naive models are fairly well scattered among the different industries. More pertinent seem to be the interrelations between relative accuracy, extent of seasonality in employment and the method apparently used by the firm to prepare the anticipations. Although the individual firms could not be contacted, it was obvious from the data that many of them were taking as their anticipation the current level of their employment, i.e., $E_{t+2}=A_t$ occasionally rounded off to the nearest 0 or 5, or generally carrying forward the previous anticipation; the code letters of these firms are italicized in Table 323. It is significant that these firms are heavily concentrated in industries characterized by little seasonal variation in employment; this is true of 19 of these 27 firms. To take a different perspective, 53 per cent of all firms in industries with little seasonal variation resorted to mechanical projection as compared with 35 per cent of firms in the other industries. Moreover, the accuracy of the anticipations varies noticeably with these two factors, as is brought out by the following tabulation, showing the proportion of firms in each category whose two-month anticipations outperform the naive model more than half the time:

	Model I	
	Mechanical projection	Other means
Substantial seasonal	50%	79%
Little seasonal	84	61
	Model II	
	Mechanical projection	Other means
Substantial seasonal	63%	79%
Little seasonal	84	72

The results for the four-month anticipations are essentially the same.

Thus, it appears that, on the one hand, firms in industries characterized by little seasonal variation were more likely to utilize simple mechanical projection in their forecasts and, on the other hand, these forecasts were substantially better than other means only in these industries. This is not unexpected, particularly since during the period studied amplitude of seasonal variation of employment was a dominant factor in total amplitude of employment.[8] Hence,

[8] R. Ferber, *Op. cit.*

mechanical projection is likely to possess high relative accuracy where seasonal variation is relatively low and is likely to yield poor results in other situations.

Overall, Table 323 indicates that the anticipations of firms relying on methods other than simple mechanical projection turn out only slightly better than those of firms relying on mechanical projection. However, when a similar analysis is carried out taking into account the magnitudes of individual errors, the margin of superiority increases substantially.[9] Thus, it appears that these other methods are most successful primarily in forestalling relatively large errors rather than in terms of the frequency of relative accuracy.

Correlation Tests. If the anticipations contain some consistent bias, the accuracy of the anticipations in gauging change can be improved by correlating the actual and anticipated *rates* of change. If little seasonal variation is present in the data, this can be done by correlating A_{t+2}/A_t with E_{t+2}/A_t for the two-month anticipations, and A_{t+4}/A_t with E_{t+4}/A_t for the four-month anticipations. However, this is not a desirable procedure if pronounced seasonal variation exists in the data, for the common influence of the seasonal on both the actual and the anticipated data tends to produce spuriously high evidence of accuracy. Hence, for the six industries and one other firm in the sample where pronounced seasonal variation was detected, the data were adjusted for seasonal (using the link-relative method) before correlating the actual and anticipated rates of change with each other.

The results of carrying out these correlations for the individual firms and the employment totals of each industry are presented in Table 327 for both sets of anticipations. The coefficients of determination between the actual and anticipated rates of change for the different firms in an industry are shown in an adjoining column but on the same line as the coefficient for the industry. As before, the coefficients for the firms are presented in order of decreasing employment, so that figures in corresponding positions of the two "individual firms" columns pertain to the same firm and are comparable with data for the same firms in Table 323.

The fundamental difference between the naive model approach used previously and the correlation test lies in the fact that in the first instance the anticipated change is used as reported, whereas the correlation test compares the actual change with the "best" linear function of the anticipations, "best" being defined in a least squares sense. It is because of this allowance by the correlation process for linear bias that the naive model tests and the correlation approach may yield different results.

The general picture presented by Table 327 is one of poor correspondence between anticipated employment and actual employment. At the industry level, a significant correlation between the two sets of data is apparent for only seven of the fourteen industries for the two-month anticipations and five of the fourteen industries for the four-month anticipations. The latter five industries are included in the former: a significant four-month correlation is accompanied by a significant two-month correlation. On the whole, the two-month anticipa-

[9] *Ibid.*

TABLE 327

CORRELATION BETWEEN ACTUAL AND ANTICIPATED RATES OF CHANGE IN LABOR FORCE, FOR INDUSTRIES AND INDIVIDUAL FIRMS

Industry	Coefficient of determination between actual and anticipated change two months ahead		Coefficient of determination between actual and anticipated change four months ahead	
	Industry	Individual firms	Industry	Individual firms
No seasonal adjustment				
Food, excl. meat packing	.31**	.38, .40**, .09, *.00*	.09	.08, .35**, .10, *.00*
Meat packing	.20	.19, .40**, *.08*	.01	.13, .42**, *.55***
Iron and steel	.24*	*.05*, *.05*, .10, .01, (−).10, .00, *.09*, *.42***, *.00*, (−).*06*, *.24**, *.25***	.17	*.02*, *.02*, .09, *.03*, (−).07, .02, *.00*, *.25**, *.00*, (−).*09*, *.28**, .29**
Other fabricated metals	.59**	.51**, .35**, *.14*	.49*	.06, .21*, *.08*
Nonelectrical machinery	.10	.42**, (−).*04*, *.59***, *.30***, *.50***, *.09*	.15	.31**, *.09*, *.69***, *.44***, *.35*** *.13*
Electrical machinery	.40**	.40**, 49**, .21*	.59**	.58**, .45**, .17*
Automobiles and parts	.26	.35**, *.00*, *.10*	.16	.26*, *.00*, *.25**
Utilities	.43**	*.17*, .61**	.63**	*.45***, .46**
Seasonally adjusted				
Printing and publishing	.01	.61**, .07, .17	.00	.64**, .08, .25**
Cans	.10	.25*, *.01*	.20	.42**, *.05*
Farm machinery	(−).10	*.00*, *.09*, *.08*	.06	*.20*, *.06*, *.04*
Communications eqt.	.42**	.27*, .45**, .25*, *.02*, .08, .08	.31*	.48**, .56**, .36**, (−).*07*, .06, .06
Other transp. eqt.	.49**	.48**, .56**, .66, *.19**, *.66***	.29*	.22*, .43**, *.01*, *.25**, .63**
Retail trade total	.13	.03, .34**, .16, .20*	(−).01	.20, .26*, .17, .69**

Note: One asterisk indicates significance at the .05 probability level; two, at the .01 level.
Italicised figures represent firms using simple mechanical projection.
Negative signs in parentheses preceding a coefficient indicate negative correlation.

tions gauge the actual change in employment better than the four-month anticipations, as one would expect.[10]

The ability of the individual firms to anticipate correctly the change in actual employment varies widely. Roughly half of the firms in the sample were able to anticipate the change in their employment more closely than would be anticipated as a result of chance. Most of these firms, however, were in one of five industries—other fabricated metals, nonelectrical machinery, electrical machinery, communications equipment, and other transportation equipment With the exception of communications equipment, the anticipations of more than half the firms in each of these industries were correlated significantly with actual changes in employment, for both the two-month and the four-month anticipations. Interestingly enough, each of these industries turns out producers' durable goods.

[10] An alternate approach, suggested by a reviewer, is to incorporate a seasonal variable in the correlation, computing the regression of the actual change in employment on the anticipated change and on the ratio of seasonal coefficients for the months in question. This procedure makes possible an assessment whether the anticipations have any *net* predictive value over and above the seasonal factor. Exploratory calculations using this technique serve to reinforce the inferences drawn from Table 3. The net correlations between the actual and anticipated changes are somewhat lower than the gross changes, particularly for firms and industries with pronounced seasonal variation, but generally remain statistically significant if they originally were so.

As a general rule, firms that are able to anticipate changes in employment better than chance two months ahead are able to do the same four-months ahead. In fact, slightly more firms were able to anticipate successfully on this basis four months ahead than two months ahead! Of the 59 firms shown in Table 327, 30 sets of four-month anticipations were correlated significantly with actual changes in employment as compared with 28 sets of two-month anticipations; 26 firms were successful on both bases. In most cases, however, the closeness of the relationship is not at all high.

On the whole, significant correlations on an industry basis tended to occur where the actual and anticipated rates of change were correlated significantly for half or more of the firms in the industry. Mathematically, there need be no relationship between these two factors, and in fact, Table 327 reveals some notable exceptions. Thus, the two-month anticipations of only three of the twelve iron and steel firms—and all three were relatively small, were significantly related to actual change but the correlation between the aggregated magnitudes was statistically significant. On the other hand, although four of the six nonelectrical machinery producers were able to anticipate changes in their employment better than chance four months ahead as well as two months ahead, the industry magnitudes showed no correspondence on either base.[11]

Easily the most important distinguishing characteristic of the successful firms was the method used to prepare the anticipations. Of the firms using one of the two forms of mechanical projection, only 5 of 27, or 19 per cent, anticipated two-month changes in employment better than chance whereas of the 32 firms using some other method, 23, or 72 per cent, were "successful" in this respect. In contrast to the accuracy of the level of the anticipations, the method used is clearly a fairly sensitive indicator of the ability of the firms to anticipate rates of change.

Direction and Turning Point Tests. Considerable care must be exercised in comparing direction of change lest the results be influenced by the same serial correlation effects we are trying to avoid. A purely mechanical approach of projecting into the future the same direction of movement in employment as occurred in the recent past would have been correct, in this study, more than half of the time purely because of this serial correlation effect.[12]

To abstract from this difficulty, comparisons are shown in Table 329 not only on the basis of ability to anticipate correctly increases or decreases in employment but also on the basis of ability to anticipate correctly *reversals* in employment. Table 329 presents four statistics for each firm in the sample as well as for the cumulated industry anticipations. These are the proportion of times an increase in employment was correctly anticipated, the proportion of declines correctly anticipated, and also the corresponding percentages for reversals in employment—upturns and downturns. Reversals are defined as changes in the direction of movement of employment from the most recent movement, after adjustment for seasonal if any, and are therefore included in the first two com-

[11] It might be noted at this point that substantial differences were apparent between industries in the average variability of actual employment from month-to-month, even after seasonal adjustment—the range for industries being from less than one per cent for utilities to over 6 per cent for meat packing. However, the correlations in Table 3 showed little overall relationship to these differences.

[12] Based on tabulations of the length of runs in the direction of change for the 14 industry totals. The actual percentages varied between 45 and 70%, with most between 50 and 60%.

TABLE 329

ACCURACY OF TWO-MONTH ANTICIPATIONS IN GAUGING THE DIRECTION OF MOVEMENT OF EMPLOYMENT, BY FIRM AND BY INDUSTRY

Industry	Firm[a]	Proportion of times direction anticipated correctly on		Proportion of times reversals gauged correctly on	
		Increases	Decreases	Upturns	Downturns
Food, excluding meat packing	A1	82*%	82*%	86%	86%
	A2	67	69	100*†	75†
	A3	100	0	100*	.00
	A4	0	0	.00	.00
	Industry	91*	73	100*	80
Meat packing	B1	50	83*	60	40
	B2	100*	14	100*	20
	B3	15	22	29	.50†
	Industry	.50†	70	60	60
Printing and Publishing	C1	50	64	100*	100*
	C2	83	56	67	57
	C3	75†	83	50	100*
	Industry	50†	50†	100*	57
Iron and steel	*D1*	67	62	100*†	0
	D2	31	0	75†	0
	D3	90*	67	100*†	50
	D4	53	13	50†	0
	D5	62	58	100*†	75
	D6	80*	10	86*	14
	D7	77*	16	80	25
	D8	70	18	60	33
	D9	85*	25	67	0
	D10	0	0	0	0
	D11	33	0	0†	0†
	D12	65	40	50	40
	Industry	92*	33	100*	.50†
Cans	E1	67	17	60	40
	E2	40	43	67†	50†
	Industry	60	25†	80	40
Other fabricated metals	F1	83*	71	80	75†
	F2	85*	11	75†	0†
	F3	67	0	100*†	0†
	Industry	86*	40	100*†	0†

* Significantly above .50 at .05 probability level.
† Based on less than five observations.
[a] Italics indicate firms using mechanical projection.

TABLE 329—(*continued*)

Industry	Firm*	Proportion of times direction anticipated correctly on		Proportion of times reversals gauged correctly on	
		Increases	Decreases	Upturns	Downturns
Nonelectrical machinery	G1	100*	08	100*†	0†
	G2	0†	50	50†	100†
	G3	94*	83	67†	67†
	G4	86	80*	50†	50†
	G5	67	50†	0†	50†
	G6	75	29	25†	25†
	Industry	100*†	29	100*	0†
Farm machinery	*H1*	0†	67	33†	50†
	H2	0†	33	0†	100*
	H3	33†	70	67†	80
	Industry	0†	57	67†	75†
Electrical machinery	I1	100*	64	100*†	67
	I2	100*	11	100*	0†
	I3	100*	17	100*	25†
	Industry	100*	63	50†	33†
Communications equipment	J1	71	17	33†	33†
	J2	88*	50	100*†	50†
	J3	78*	43	100*†	50†
	J4	100*	82*	100*†	67†
	J5	75†	36	50†	40
	J6	100*	38	100*†	75
	Industry	100*	33	67†	33†
Autos and parts	K1	100*	0	100*	0
	K2	0	10	0	17
	K3	75	12	50†	0†
	Industry	100	9	100†	0
Other transportation equipment	L1	100*	63	0†	50†
	L2	67	29	100*†	100*†
	L3	80	20	75	14
	L4	75†	57	67†	50†
	L5	82*	89*	67†	33†
	Industry	63	0†	100*†	0†
Utilities	*M1*	28	9	33†	25†
	M2	100*	42	100*	50†
	Industry	71	25	67†	33†
Retail trade	N1	100†	0†	40†	67
	N2	0†	67	50	75†
	N3	57	50	60	100*†
	N4	50	80	100*	71
	Industry	0†	67†	100*	100*

* Significantly above .50 at .05 probability level.
† Based on less than five observations.
* Italics indicate firms using mechanical projection.

parisons as well. As in Tables 323 and 327, the firms in each industry are listed in order of decreasing size. As before, code letters of firms relying on simple mechanical projection are italicized.

The asterisks in the table indicate which of these percentages significantly exceeds 50 per cent at the .05 probability level or beyond, thus indicating the extent to which the anticipations are superior to the accuracy that might have been obtained as a result of chance. It is important to stress, however, that this interpretation is subject to qualification with regard to the anticipation of increases and decreases because of the serial nature of these phenomena.

In an aggregate sense, the results in Table 329 can be summarized in the following three lines:

	Increase	Decrease	Upturn	Downturn
Number of firms with "success" ratios significantly above .50	21	5	20	5
Per cent of all firms	35%	8%	34%	8%
Number of industries with "success" ratios significantly above .50	6	0	7	1

It is clear from these data that at best only one-third of the firms were successful in anticipating increases or upturns in employment better than what would be expected as a result of chance, and that only a small fraction of the firms were successful in this regard when it came to declines or downturns in employment. At the same time, when the firm anticipations are aggregated, roughly half of the industry anticipations successfully predict increases or upturns in employment better than chance but almost none of the industry anticipations did so when employment moved in the other direction.

Table 329 indicates that in general a negative correlation exists between a firm's ability to anticipate upward movements and its ability to anticipate downward movements. Thus, 23 firms were able to anticipate *either* increases *or* decreases in employment better than chance, but only three were successful in both respects! The same is true of upturns and downturns. This rather remarkable characteristic would seem to indicate that the firms with the more successful anticipations in this regard are not necessarily the more astute but rather are those forecasting most consistently in a particular direction! If upward movements are combined with downward movements, we find that only three of 59 firms were able to anticipate the direction of change significantly better than chance.

As before, the method of anticipation was important, as shown by the following tabulation:

	Increase	Decrease	Upturn	Downturn
Proportion of firms using mechanical projection and anticipating direction better than chance	4/27	3/27	3/27	1/27
Proportion of firms using other methods and anticipating direction better than chance	17/32	2/32	18/32	4/32

STRUCTURE OF THE ANTICIPATIONS

The extent to which anticipations can in turn be anticipated is a central issue in any study of such data, for if they can be anticipated with a high degree of precision their net value for analytical purposes is likely to be very small. In at least one study, this has proved to be the case.[13] An important corollary question which has hardly been touched, however, is the extent to which hypotheses explaining industry anticipations are equally successful in explaining the anticipations of the individual firms that comprise the aggregate. Does the hypothesis explaining industry anticipations typify individual firms as well, or is it in effect some hybrid phenomenon representing a conglomeration of diverse structures of firm anticipations but representing none of them?

In the analysis of the Illinois employer's labor force anticipations, six different hypotheses were advanced to explain the structure of the anticipations, as shown in Table 333.[14] Each of these hypotheses was applied to the anticipations of the appropriate firm and industry total in turn, according to the seasonal amplitude of the data, by means of a series of reliability tests ending in multiple regression analysis.[15] The results indicated the extent to which each hypothesis was successful in explaining the structure of the anticipations of particular firms and industries. We are therefore able to compare the extent to which a hypothesis successfully explaining industry anticipations is equally successful for the individual firms in the industry. A summary comparison of this type is shown in Table 334.

Two findings relative to the diverse nature of the results are evident in Table 334. One is that considerable variation exists with regard to which hypothesis best explains the structure of a particular industry's anticipations, at least of those considered here. The only hypothesis which appears with any regularity is F, the mechanical extrapolation hypothesis. This result may not be unexpected considering the substantial differences in seasonal variation between employment in different industries.

The other finding, of more relevance to the subject at hand, is that a hypothesis which successfully explains the structure of anticipations of a particular industry seems as likely as not to represent the structure of anticipations

[13] Ferber, Robert, *The Railroad Shippers' Forecasts*, Chapters 4–5. This is especially so if lack of predictive value is also demonstrated for the residuals from the regression of the anticipations on past experience, as was the case with the railroad shippers' forecasts.

[14] A discussion of the theoretical basis for each of the hypotheses is contained in the forthcoming monograph on this subject to be published by the Bureau of Economic and Business Research of the University of Illinois. It might be noted that these hypotheses cover a wider range of possibilities than may be evident at first sight, for many of the more complex hypotheses are but extensions of one of these. If the simpler form is not supported by statistical tests, more complex forms can also generally be safely disregarded.

[15] To reduce the amount of calculation, only those hypotheses passing certain "elimination tests" were subjected to regression analysis for each firm and industry. Thus, for (A) to reach the multiple regression stage, it had to satisfy the following two conditions: (1) the correlation between $E_{t+3}-A_t$ and A_t-A_{t-3} should be statistically significant at the .05 probability level; (2) the value of r^2 in the preceding case should exceed the corresponding correlation between $A_{t+3}-A_t$ and A_t-A_{t-3} by at least .05. The latter condtition is somewhat arbitrary, necessitated by the failure of the usual significance tests to discriminate between the two correlations, apparently because of the small sample sizes involved. For each firm and industry total in which hypothesis A satisfied these two conditions, the operational form of A was estimated by least squares using E_{t+3} and A_{t+3} as alternate dependent variables. If the least square form of (1) with E_{t+3} dependent survived certain additional tests—primarily similar forms of the above two and that the estimate of c be statistically significant while the estimate of b not differ significantly from 1.0—the hypothesis was accepted for that particular firm or industry. Note that if any of the first five hypotheses satisfy all the requirements but c is not significantly different from zero, we are automatically left with hypothesis F. A discussion of these tests is included in the publication mentioned in the preceding footnote.

of a minority of firms in the industry. In only four of the 13 instances in which a hypothesis successfully explained the industry structure of anticipations was it equally successful in explaining the structure of anticipations of more than half of the firms in the industry (though in one other industry just half were explained—one out of two). Furthermore, the degree to which a hypothesis is successful in explaining industry structure, as measured by the coefficient of determination, does not seem to provide much indication of the extent to which the same hypothesis explains the anticipations of the firms in the industry. Thus, in three of the six instances in which a hypothesis explained more than 90 per cent of the variation in industry anticipations the same hypothesis was acceptable for the anticipations of a minority of the firms in the industry; and

TABLE 333

ALTERNATIVE HYPOTHESES ON STRUCTURE OF ANTICIPATIONS

Amplitude of seasonal variations*	Hypothesis†	Operational form†
Slight	(A) $A_t+(A_t-A_{t-2})$	$a+bA_t+cA_{t-2}$
Slight	(B) $A_t\left(1+\frac{A_t-A_{t-2}}{A_{t-2}}\right)$	$a+bA_t+c\frac{A_t^2}{A_{t-2}}$
Substantial	(C) $A_t'+(A_t'-A_t'_{-2})$	$a+bA_t'+cA_{t-2}'$
Substantial	(D) $A_t'\left(1+\frac{A_t'-A_t'_{-2}}{A_{t-2}'}\right)$	$a+bA_t'+c\frac{A_t'^2}{A_t'_{-2}}$
Both	(E) $A_{t-10}\left(1+\frac{A_t-A_{t-12}}{A_{t-12}}\right)$	$a+bA_{t-10}+cA_{t-10}\frac{A_t}{A_{t-12}}$
Both	(F) A_t	$a+bA_t$

* Based on the derivation of seasonal indices by the link-relative method, the boundary between "slight" and "substantial" amplitude being 10 per cent variation from trough to peak.

† Dependent variable in all instances is E_{t+2}.

Notation: A_{t-n} denotes actual labor force n months ago.
E_{t+2} denotes anticipation two months ahead, prepared in month t.
Primes denote seasonally-adjusted values.

the hypothesis with the lowest industry coefficient of determination (iron and steel) proved successful on seven of the 12 firms in the industry, producing coefficients of determination for those seven firms ranging between .90 and .998.[16]

The indifferent success of the acceptable industry hypotheses with the individual firm anticipations points to the very real possibility that industry anticipations represent at times a conglomeration of diverse structures of individual firm anticipations, as noted in the opening paragraph of this section. This is borne out by further examination of Table 334, which reveals considerable disparity in the structure of the anticipations of the different firms in the same industry; indeed, in this respect there is at least as much variability within most industries as between them. Whether or not a particular hypothesis is

[16] The three instances where a successful industry hypothesis also explains the anticipations of most or all of the firms in the industry all involve the simplest hypothesis, F. Therefore, the implication of the fact that the estimates of the slope, b, for the firms are close to the corresponding estimate of b for the industry remains a moot point.

TABLE 334

SUCCESS OF HYPOTHESES IN EXPLAINING THE STRUCTURE OF THE ANTICIPATIONS OF INDUSTRIES AND FIRMS

Industry	Hypothesis acceptable for industry total			Other hypotheses acceptable for firms in given industry	
	Hypothesis	R^2	Proportion of firms in industry for which hypothesis holds	Hypothesis	Number of firms
Food, excl. meat packing	A	.77	1 of 4	F	1 of 4
	B	.77	0 of 4		
Meat packing	None		—	F	1 of 3
Printing and publishing	None		—	D	1 of 3
Iron and steel	F	.69	7 of 12		—
Cans	None		—	F	1 of 2
Other fabricated metals	F	.93	1 of 2		—
Nonelectrical machinery	F	.98	5 of 5		—
Farm machinery	F	.997	3 of 3		—
Electrical machinery	F	.87	1 of 3		—
Communications eqt.	F	.84	1 of 5	C	1 of 5
				E	2 of 5
Autos and parts	F	.98	2 of 3		—
	E	.88	1 of 3		
Other transp. equipment	E	.94	1 of 5		—
	F	.94	1 of 5		
Utilities	None		—	F	1 of 2
Retail trade	D	.86	0 of 4		—

applicable to both industries and firm anticipations cannot be deduced from results pertaining to the former alone, even apparently if well over 90 per cent of the variation in industry anticipations is explained by that hypothesis.

IMPLICATIONS OF RESULTS

This study reveals considerable disparities between the nature of results obtained at the individual firm and industry levels with regard to the accuracy and structure of businessmen's anticipations. The anticipations are clearly more accurate at the industry level, a superiority which is attributable in the main not to equally high accuracy of individual firm anticipations but rather to compensating errors of anticipation by the individual firms. Similar results have been obtained in the analysis of capital expenditure plans.[17]

Also of considerable interest is the finding that very few firms were able to forecast employment consistently better than chance *in both directions*. Firms with the most accurate prediction records were generally those which predicted changes consistently in the direction which coincided with the trend of their employment (generally upward); a change in employment counter to this trend was invariably missed. At the same time there is evidence that firms

[17] See papers by Keesor and by Foss and Natrella in forthcoming proceedings of Conference on Anticipatory Data, held by the Universities-National Bureau Committee at Princeton, Nov. 8-9, 1957.

which make some effort to anticipate employment, i.e., firms that do not consistently mechanically extrapolate current levels, are more accurate than other firms on the basis of tests of relative accuracy.[18] In an absolute sense, such firms do show larger errors of the forecasts, but only because of a tendency for mechanical extrapolation to be used more frequently in industries characterized by relatively low fluctuations in employment during the period studied.

This study also provides corroborative evidence for a hypothesis advanced in connection with the railroad shippers' forecasts to the effect that the structure of individual firm anticipations may be wholly dissimilar to that of the aggregate; and that the latter represents an average of two opposite types of individual anticipations—mechanical extrapolation of level and relatively accurate non-mechanical forecasts.[19] The findings of this study point to the existence of several, rather than two, different types of forecasts insofar as structure is concerned, with mechanical extrapolation of level as the predominant type. With a few exceptions, these different types intermingle in each industry so that the industry-derived structure of the anticipations is likely to bear little relation to that of most of the individual firms. As the level of aggregation rises to a regional and national scale, the dissimilarity between aggregate and individual-firm structure will undoubtedly increase.

From a more general point of view, this study serves to highlight a basic difficulty of these and other types of anticipations solicited from the outside. If the particular data requested are germane to the firm's plans and operations, serious attempts at accurate forecasts (resulting in, apparently, more successful forecasts) can be expected. If the data appear to be external to the firm's operations, there is likely to be an inherent inclination to comply with the request with the least effort. The problem of obtaining useful expectations data, therefore, entails not only the construction of a meaningful set of concepts from an analytical point of view but also of fitting these concepts into the operational framework of the firm.

[18] This is corroborated by a similar study of the Indiana labor force anticipations supplemented by personal interviews conducted by Professor William Andrews of Indiana University. Firms that were considered to be making some effort to anticipate employment were generally much more accurate than other firms.

[19] Hearings Before the Subcommittee on Economic Statistics of the Joint Committee on the Economic Report, *Report of Federal Reserve Consultant Committee on General Business Expectations*, 1955, pp. 570–571.

AN INDEX OF MANUFACTURING PRODUCTION IN NEW ENGLAND*

HARRY BENJAMIN ERNST
Tufts University

Part of the current production estimates of the Federal Reserve Board are based on projected and interpolated productivity relationships which are two or three years old. An alternative method is outlined which is in three parts.

1. Annual relationships are estimated for kilowatt hours, production worker man-hours and output, based on plant data obtained by mail survey.
2. Man-hours and KWH are collected by the 20th of each month for specific two and three digit Standard Industrial Classification Industries.
3. The annual productivity relationships are applied to the monthly man-hours and KWH in making current, monthly production estimates for specific industries.

There is still a lag in obtaining annual relationships, but the monthly KWH and man-hour data are helpful indicators of "scale" or "technological" changes which affect productivity. Illustrative but preliminary total manufacturing indices for New England are presented.

In any highly industrialized region or economy, current measures of manufacturing output are of great value. Manufacturing output is significant both because it is a large part of total output of goods and services and because of its interrelations with other sectors of the economy. Until 1955, employment data were the chief source of manufacturing information in New England. Except for very short time periods, total employment data are unsatisfactory indicators of industrial output because output also depends on the number of hours worked and on the technological and organizational structure of the industries involved. It is true that an activity index could have been assembled from the Bureau of the Census, "Boot and Shoe Production," "Cotton Spindle Activity" or other series for New England. However, there are only a few such indicators and no broad regional measure of output comparable to the Federal Reserve Board's *Index of Industrial Production*. In view of this, the Federal Reserve Bank of Boston has been financing research relating to detailed measures of the region's manufacturing output. The basic research is being carried out at the Institute for Economic Research, Tufts University.

* This paper explains procedures used in constructing the New England Indices of Manufacturing Production at the Institute for Economic Research, Tufts University, under a research grant from the Federal Reserve Bank of Boston. The paper describes the problems and general methodology of the research. Only preliminary, very crude results are presented. The indices are being prepared for the New England region as a unit, and for the states of Maine and Massachusetts separately. Much credit should go to staff members Daniel Ounjian, Gerald B. Dahill, Allan Gard and Ira Stepanian; also to Lewis F. Manly of Tufts University; George H. Ellis, Director of Research, Federal Reserve Bank of Boston; Wendell D. MacDonald, Director, and Edward T. O'Donnell of the Boston Office, Bureau of Labor Statistics; Joseph F. King, Director, Massachusetts Department of Labor and Industries; R. Leigh Fitzgerald of the New England Electric System and William J. Ray of the Electric Council of New England, as well as others too numerous to mention.

Since the FRB model is closest to the new model, a brief description of some aspects of the FRB methods follows. About fifty per cent of the FRB monthly output indexes are estimated from changes in the production worker man-hours reported by the U.S.Bureau of Labor Statistics. Chart 338, reproduced with the permission of the Federal Reserve Board, Division of Research and Statistics, illustrates the method for a hypothetical industry. Essentially, the relationship between output and man-hours is determined annually, and then nterpolated in the form of a smooth curve within given years. Extrapolation is also necessary since annual data lag current monthly man-hour inputs by two or three years. Annual relationships are applied to monthly variables in both the interpolative and extrapolative procedures. One must assume that the relationships between output and man-hours change smoothly within years and the relationships may be extrapolated smoothly through time.[1] As an alternative to the FBR procedures, there is the model described below.

I. THE CONCEPTUAL MODEL FOR THE NEW INDEX OF MANUFACTURING PRODUCTION

As with the FRB index, annual-monthly relationships are used in the new model, but with essential differences in the basic units of measurement. Monthly production worker man-hours are supplemented with kilowatt hours for specific industries. Annual man-hour and kilowatt hour input data are used plus deflated value of production, (because physical output data are not available either monthly or annually). The two monthly input series are compiled for two-digit SIC classifications (and in some cases, three-digits) for Massachusetts, Maine and the remainder of New England.

From annual plant, man-hour, kilowatt hour and deflated VOP data, production functions are constructed for each important industry or subindustry. The basic form of the production function takes into account the factors of "technological" and "scale" changes.[2] As a working hypothesis, if the level of technology is known (as indicated by a KWH/MH ratio in those industries in which KWH are an important input), one has an essential, identifying characteristic of a production function. Once having specified a production function, the scale effect (increasing or decreasing returns) can be separately determined.

Two situations are included in our concept of "scale," which are part of the working hypothesis. First, in some industries, there might be changes in the volume of production because of changes in the number of shifts worked, with no changes in the stock of capital equipment itself. On the assumption of constant capital stock, changes in output are a matter of the degree of utilization of the capital equipment and any change in labor productivity

---

[1] In a paper presented at the Annual Meeting of the American Statistical Association at Atlantic City on September 11, 1957, Milton Moss of the Federal Reserve Board, Research Division, presented tests of the interpolative assumption in those industries for which physical output and man-hours were available monthly. Estimates of production resulting from a smooth interpolation (within years) of annual output and man-hour relationships did not differ significantly from the actual physical output. However, these results pertained to aggregates of industries and not specific industries, and only to the interpolative procedure.

[2] Technological and scale effects have been mentioned by many writers. For example, G. W. Wilson, "An Influence of Demand Upon Productivity," Canadian Journal of Economics and Political Science, Aug. 1956, Vol. XXII, No. 3; p. 379; and Witt Bowden "The Productivity of Labor in Great Britain," *Journal of Political Economy*, June 1937, Vol. 45, No. 3, p. 347.

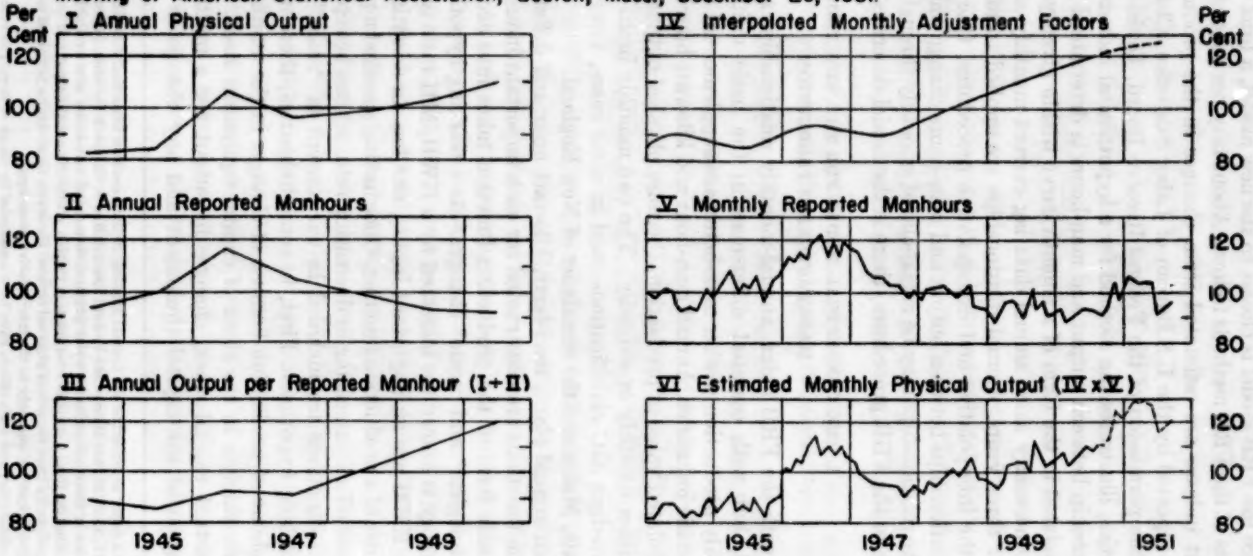
CHART 338

USE OF MANHOUR DATA FOR INTERPOLATION IN FEDERAL RESERVE INDEX

All Indexes, 1947–49 average = 100

SOURCE: "Progress in the Revision of Federal Reserve Index of Industrial Production" *

* Paper prepared by Lorman C. Trueblood and Arthur L. Broida, Division of Research and Statistics, Board of Governors of the Federal Reserve System, for session on "Index Number Developments," annual meeting of American Statistical Association, Boston, Mass., December 29, 1951.



should properly be associated with the "volume effect." In other industries or time periods, there could be increases in the stock of capital equipment, as an extension of the capital stock already employed (which is more consistent with the concept of "scale" changes). A special case of "scale" changes would be to decrease plant or industry capacity by failing to replace capital stock as it wears out. For "volume" as against "scale" changes, one would expect to find different relationships, at the plant level, between KWHs, man-hours and output.[3] If working with sufficiently disaggregated data, one would expect to find similar relationships at the industry level due to competitive plants imitating one another's technology. In cases of "volume" (but not "scale") changes, the functional relationship of KWH and man-hours ought to have a positive intercept because of the fixed elements in KWH which are not directly associated with production changes (e.g. air conditioning, lighting, etc.) "Scale" changes ought to result in nearly proportional KWH and man-hour changes over time and more nearly homogeneous functions.

On the other hand, "technological change" is defined to include the introduction of a new kind of capital equipment to the production process. To illustrate, if a woolen and worsted broadcloth plant expanded capacity by buying additional mules, this would be a "scale" change. However, a shift from mules to more modern spinning frames would be a technological change. Typically, technological changes result in a much larger amount of electric power consumption and a new KWH and man-hour relationship. Technological changes, therefore, ought to be associated with shifts in the KWH/man-hour functions for plants and industries—but not necessarily with changes in the KWH/man-hour ratios, due to the non-homogeneous nature of the KWH/man-hour function.

There are two basic forms of production function, either of which might be chosen on an empirical basis, industry by industry.

1. Output/MH = fKWH/MH, MH) which is a productivity function. It is hypothesized that the relationship of the independent variables will indicate whether "volume," "scale," or "technological" changes are influencing labor productivity.
2. Output = f(KWH, MH) which is the usual production function.[4] In this case, the KWH/man-hour ratio could be used as the criterion of "technology" or "scale" changes and either or both the KWH and man-hour inputs used to estimate the "volume" effect.[5]

Having selected the form of production (or productivity) functions on the basis of the annual plant data, the annual relationships are used together with monthly input indexes to make monthly estimates of production changes. In a given month, production is estimated from man-hours (like the FRB) but also

[3] Empirically, on the basis of KWH and man-hour data for specific industries, one may be able to separate "volume," "scale," and "technological" changes, each having a different influence on labor productivity.

[4] E.g. P. H. Douglas, *The Theory of Wages* (New York 1934) pp. 209–11.

[5] Two sets of scatter diagrams of plant data have provided the basis for selecting the first or second type of function. On the first set of diagrams, output/man-hour was regressed on man-hours and then KWH/man-hours. On the second set of diagrams, output was regressed on man-hours and then kilowatt hours. The (multiple) correlation coefficients are the final criteria for selecting either type of function, based on linear (multiple) regression functions.

from the additional input of KWH (which of course, affects the level of production and labor productivity). It is expected that the addition of monthly KWH data will permit more accurate interpolation and especially extrapolation of output/man-hour relationships, even though the application of annual relationships to monthly data still entails some smoothing of the output estimates.

II. CONCEPTUAL PROBLEMS ASSOCIATED WITH THE NEW MODEL

First, in using both man-hours and KWH in the same function, one must recognize a certain amount of non-comparability in the two series. There are some well-known limitations to using production worker man-hours to reflect month to month production changes. The Bureau of Labor Statistics man-hours are those paid for but not necessarily worked, causing difficulties in vacation periods and months during which major holidays occur. The man-hours also cover the payroll period (usually a week) in the middle of the month —which might not always be representative of the activity during the other three or three and one half weeks of the month. On the other hand, KWH data reflect activity for the entire month (and the actual activity, in fact, since KWH are not paid for unless used). Furthermore, there are probably fewer discontinuities in output/KWH than in output/man-hour relationships—for example, output/man-hour relationships can change during slack periods due to management's desire to keep the work force intact. In most industries, month to month changes in activity are probably more accurately reflected by KWH than by man-hours. This of course, is not true in the shoe, paper box or other industries where power consumption per man-hour is relatively insignificant. Furthermore, because of the non-comparability of the two inputs, one must be careful in assessing short run shifts in KWH/man-hour ratios.

The choice of data transformation and production function model creates a second kind of problem. At least three basic models could be used. The first might utilize a purely cross-sectional approach—for example, regressing output on both inputs, with the data in raw form. This method is not satisfactory because of wide differences in the average productivity of plants. Shifts in the importance of plants (some growing, some declining) produce a "shot-gun" effect on a scatter diagram—aside from any changes in the plant input-output relationships themselves. As another objection, if large and small plants had different input-output relationships, a function of this type would reflect these plant-size differences, since the large values underlying the function would represent only the relationships of the large plants and the smaller values the relationships of the small plants.

A second function model might be based on changes in the cross-sectional data, with yearly changes in the input-output data correlated. This model could be used when it is desirable to reduce the auto-correlation in the input series in order to estimate a behavioral parameter for each input. Since first differences of the man-hour and KWH inputs do not take account of plant size differences and probably aren't comparable because of measurement difficulties, this approach has not been used in the present context. Furthermore, in

thinking about production functions in the light of the fixed factor analysis above, there is probably little reason to expect the KWH and man-hours to change inversely—excepting in periods of technological change. All other changes may be due to random forces or non-comparability, and short term changes in output are probably better measured by changes in the KWH (or man-hour) data alone.

The model finally selected (since it seems to fit the data best) is based on the hypothesis that percentage changes in plant output are a function of percentage changes in man-hours and KWHs. In using the two models above, one assumes that absolute output changes are a function of absolute man-hour and KWH changes (or their differences). Actually, any of the three assumptions might be the most useful under given circumstances. By using index numbers to the base period 1950–52 in the percentage change model, the auto-correlation of the three variables is probably increased (with respect to the first difference model) but the standard error of estimate is also probably reduced, which should be helpful in the extrapolative procedure. The percentage change approach was used in both the production and productivity functions already described.

A final problem concerns the reversibility and durability of the annual production functions. In fact, is a production function reversible? For example, if an industry of relatively homogeneous plants were expanding its output (with a given level of capital stock), there would result a certain pattern of input productivity relationships. Now, if the industry reduces output, need the same production function apply? What if the more effective man-power is retained or waste is otherwise eliminated, or there are serious inter-plant productivity differences, etc? The function is reversible only if the effects of other important factors underlying input productivity relationships are either held constant or cancel out. Aggregative data are not very useful in these matters. What is needed are detailed plant studies dealing with the effects of "technological," "scale," and "volume" changes, as well as other important factors underlying productivity changes.[6]

Another important aspect of the same problem is the durability of the production function to be used for extrapolation. As one approach, we have suggested investigations of the relationships between "technological" or "scale" changes and KWH/man-hour functional relationships. It is probable that plant and equipment expenditure surveys may reveal important changes to come, which can be verified and timed by KWH/man-hour changes. Inquiries regarding "technological" or "scale" changes can be directed to key plants. However, "techniological" or "scale" changes are not simple matters. All plants in an industry do not change their technology or scale simultaneously. The deepening or widening of capital (as measured by KWH/man-hour ratios) begins with one or only a few plants and there may be no noticeable effect at the industry level. As the new processes or increases in capacity are diffused,

[6] As one example of detailed plant studies, see S. Melman, *Dynamic Factors in Industrial Productivity*, John Wiley & Sons Inc. New York, 1956.

particularly to the larger plants, a new industry production function has evolved.

The above problems are not exhaustive but merely point out some of the difficulties which must be dealt with in the area of aggregative industrial production and productivity measurement.

III. SOURCES OF DATA FOR THE CURRENT MODEL

To obtain the annual plant data for the production functions, a mail sample of about 1200 New England companies was asked to report the annual production worker man-hours, KWHs, value of production (not sales) and physical output wherever possible.[7] It proved difficult to deflate the value of production data. For the most part the commodity indexes of the Wholesale Price Index were used, supplemented by implicit price indexes derived from plant physical output and value of output data. Each plant was then coded by its principal product in accordance with the Standard Industrial Classification, using the U. S. Bureau of Labor Statistics interpretation of the code wherever doubt arose as to the classification of large, multi-product plants. From these annual plant data were derived the annual production functions, for the interpolative and extrapolative procedures.

Neither of the monthly inputs, production worker man-hours and kilowatts hours, were originally available. The BLS New England Regional Office advanced an intended program and began to report monthly man-hours, by two-digit SIC groups from January 1950. In Massachusetts greater detail was possible, through the Massachusetts Department of Labor and Industries.

The original obstacles were even greater in the case of monthly KWH data. First, it was necessary to enlist the support of 19 major electric utilities of the region and then to supervise the coding (by SIC) of their customers to make them comparable to the BLS man-hours.[8] Then, the reporting mechanism had to be activated by which the monthly data would be made available around the twentieth of the month. Today, all utility company reports are collected by the Electric Council of New England and the sample coverage ranges from sixty to eighty-five per cent of the two-digit SIC groups. The monthly KWH are corrected for working day differences, and seven standard holidays, using the working days published by the FRB,[9] since the man-hours are not subject to the same calendar variations from month to month. The working day correction is intended to increase the comparability of the monthly man-hour and KWH indexes.

---

[7] Approximately 600 firms (covering 700 plants) responded to this survey with usable data. These 700 plants produced about one-third of the total value of manufacturing production for New England, in the base period of 1950–52. (No BLS man-hour data were available prior to 1950, or 1947–49 would have been selected as the base period).

[8] About 125 important self-generators of electric power also report their generation to the Federal Reserve Bank of Boston for each calendar month. These data are added to the reported sales of the electric utilities, in each industry.

[9] The working days for each four-digit SIC industry were published in a reprint from the Federal Reserve Bulletin for December 1953, issued by the Board of Governors, Federal Reserve System, Washington D.C., pp. 42–53.

IV. PRELIMINARY PRODUCTION FUNCTIONS AND INDICES

The preliminary production functions were very crude, being based on the the annual aggregation of the three variables for all plants in a given industry.[10] Aggregating variables across plants for the five individual years provides five observations for each variable, and relatives (1950–52 = 100) were computed for the fifteen aggregates.[11] In case the productivity function was used (i.e. DVOP/MH = f KWH/MH + MH), the aggregated variables were transformed into KWH/man-hours, deflated value of production/man-hour and man-hour indexes.

From the aggregated relatives, usually a linear function was computed for the four years 1950–53 and a prediction made for 1954, using the 1954 input relatives and the four year function. Not much significance could be attached to either regression coefficients or results due to the great restrictions on the degrees of freedom. "Good fits" could be expected (from fitting three-variable regression functions to five observations), but little information on whether a shift in the production function had taken place. A sample illustration of this procedure follows:

TABLE 343

COMPARISON OF FOUR AND FIVE YEAR FUNCTION PREDICTIONS, FOR THE NEW ENGLAND CHEMICAL AND PETROLEUM INDUSTRY
(1950–52 = 100)

Year	KWH/MH	MH	DVOP/MH	Predictions	
				4 yr. Function	5 yr. Function
1950	97.0	94.0	97.7	99.3	97.7
51	95.8	103.8	96.1	98.7	96.8
52	107.0	102.2	106.5	106.7	106.8
53	118.5	105.3	118.8	115.0	117.1
54	132.0	101.7	128.1	124.6[P]	129.0

1950–53 (4 year) Function DVOP/MH = 27.6 + .717 KWH/MH + .023 MH.
1950–54 (5 year) Function DVOP/MH = 9.5 + .891 KWH/MH + .019 MH.
P—the value of 124.6 is obtained by substituting the 1954 inputs of KWH/MH (132.0) and MH (101.7) into the 1950–53 productivity function, as a rough test of its stability.

As expected, in the capital intensive Chemical and Petroleum Industry, productivity is more closely correlated with "technical" (rather than "volume" or "scale") changes, but with so few degrees of freedom, there was no adequate way to test the signficance of the regression coefficients. Having constructed the crude preliminary functions, as described above, the monthly KWH and MH indexes were computed relative to the base period (1950–52) and sub-

[10] However, other types of disaggregated production function models were tested, but due to different average productivities as well as different cyclical behavior of the sample plants, better results at the two-digit level were not forthcoming. Also, the scatter diagram technique described in footnote 5 on page 339 was used as the criterion for choosing either the production or productivity aggregative function for specific industries.

[11] Some 1955 data are now available and the analysis is being extended over the six year period.

stituted in the appropriate annual function. With reference to the FRB approach, the output/man-hour ratio was changed throughout a year according to changes in the monthly KWH and/or MH levels (instead of as a smooth curve between annual levels of output and man-hours). Secondly, in extrapolating output/man-hour levels one had the benefit of the KWHs (in addition to the MHs) in those years for which annual output data were not yet available. As long as the industry production functions weren't discontinuous, the KWH data were valuable additions for interpolative and extrapolative purposes.

Net Value Added by Manufacture weights (the average for 1950–52) for each two-digit SIC Industry, were used in constructing the total Monthly Index of Manufacturing Production. The weights were derived from the annual Surveys of Manufactures of the Bureau of the Census.

In Chart 345, the preliminary total New England Index of Manufacturing Production is compared with that the of FRB for the period of 1950–57. The general movements of both seasonally adjusted series are similar, with slightly more growth in the national series but a general coincidence of turning points. Annual averages are used from 1950 through 1954 and monthly data since that time, accounting for the greater variability in the series from 1955.

One check on the accuracy of the new index was provided by the net value added estimates of the Census Bureau (which were undeflated). For comparison purposes, total indexes of man-hours and kilowatt hours are included in Table 344 below. All data are for New England:

TABLE 344

COMPARISON OF THE TOTAL INDEX OF INDUSTRIAL PRODUCTION, NET VALUE ADDED, PRODUCTION WORKER MAN-HOURS AND KWH CONSUMPTION FOR NEW ENGLAND MANUFACTURING INDUSTRIES
(1950–52 = 100)

Year	Index of Manufacturing Production	Index of Undeflated Net Value Added*	Index of Total Production Worker Man-Hours	Index of Total Industrial KWH Sales & Generation
1950	91.8	89.5	96.6	96.1
51	102.6	102.1	102.8	101.7
52	107.6	108.4	100.6	102.2
53	117.5	119.3	102.5	112.0
54	105.6	109.5	89.4	108.1
55	113.9	n.a.†	92.2	119.7

* The estimates of total net value added by manufacture for New England are taken from the annual Surveys of Manufactures, Bureau of the Census, U.S. Department of Commerce.

† Not yet available.

Since the Census Net Value Added estimates were uncorrected for price changes, they could be expected to show an upward drift over the Index of Manufacturing Production. Also, the differences between the estimates of net value added and output for individual industries would tend to be larger. Yet, there is a general similarity of trends and cyclical turning points in the total

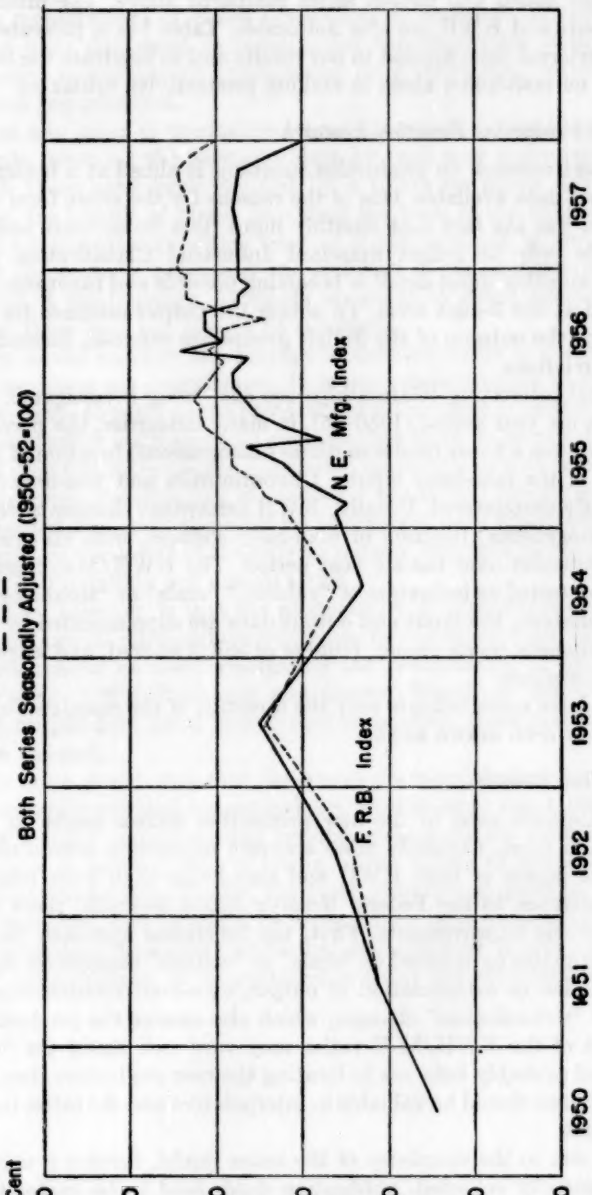


CHART 345
Comparison of F. R. B. Index with N. E. Mfg. Index
Total Manufacturing
Both Series Seasonally Adjusted (1950-52=100)
Per Cent
140
130
120
110
100
90
80
1950
1951
1952
1953
1954
1955
1956
1957
F.R.B. Index
N.E. Mfg. Index

net value added and output series compared above. The different trends of man-hours and KWH are also noticeable. Table 344 is presented as a sample of the external tests applied to our results and to illustrate the necessity of not relying on man-hours alone in making productivity estimates.

Current Production Function Research

Current research on production functions is aimed at a better utilization of the micro data available. One of the reasons for the crude form of the original function was the fact that monthly input data (man-hours and KWH) were available only in 2-digit Standard Industrial Classification groups. Now, greater monthly input detail is becoming possible and functions are being constructed at the 3-digit level. To obtain the output estimate for a 2-digit SIC industry, the outputs of the 3-digit groups are summed, instead of the input-output relations.

Several interesting relationships are also being investigated. For example, over the six year period (1950–55), in many industries, the percentage change of output was a linear (and sometimes homogeneous) function of the percentage change of the man-hour inputs. Discontinuities and non-linearities were not frequently encountered. Usually, KWH percentage changes were a linear, (but not homogeneous) function of man-hour changes, with no evidence of shifts in the function over the six year period. The KWH/Man-hour relationships are being tested as indicators of "volume," "scale" or "technological" changes. In the analysis, the input and output data are supplemented by data on plant and equipment (book value), number of shifts worked, and wages paid to production workers.

The above notes indicate only the direction of the research—no firm conclusions have been drawn as yet.

Concluding Remarks

The methods used in the new production indices ought to provide some benefits to users. Certainly more accurate production estimates can be made from the inputs of both KWH and man-hours than from man-hours alone. With reference to the Federal Reserve Board methods, these methods offer some specific improvements. First, the functional approach permits productivity estimates to be based on "scale" or "volume" changes (as against straight interpolation or extrapolation of output/man-hour relationships). Second, in cases of "technological" changes, which also change the production functions, the shift of the KWH/M-H ratios may very well signal the "technological" shift, and probably help one in locating the new production function. The first improvement should be valuable in interpolative and the latter in extrapolative procedures.

Also, due to the simplicity of the index model, involving only two components centrally collected, publication dates tend to be earlier than for more complicated indices. Both speed and coverage will be increased when all electric utilities, regardless of size, begin to use high-speed equipment for billings.

Another advantage of the new indices is that measurement is provided in areas where no measurement of either production or productivity currently exists. Also, the basic data used in constructing the production indices are valuable for studies in multi-input productivity or in forecasting future man and electric power requirements.

As is true with any index of production, there are certain limitations to the new index. Plants returning the necessary data by mail may make reporting errors (although these have been rectified by a follow-up phone call when suspected). Also, the sample of plants providing the basis of the annual functions may not be representative, either because of the exclusion of key plants which refuse to cooperate or because of the inclusion of plants which happen to be not typical of the industry. Attempts are made to guard against such possibilities, but these may not always be successful.

Since the new model employs the concept of deflated value of production, price deflation problems are considerable. It is not aways possible to find the exact commodity price to apply to the physical output or value of product data. In this event, subjective estimates and compromises have to be relied upon. It is hoped that such pricing errors are distributed randomly, so that their effect will be minor following the grouping of plants for the annual functions.

It should be noted also, that there are no benchmark data which can serve as absolute checks on the levels of the various indices. The lack of benchmark data extends to both monthly and annual series.

A key limitation, however, is that the productivity or production estimates must be based on annual relationships which are two or three years old. This limitation is one reason we are investigating the relationships of plant and equipment spending, KWH/Man-hour and Output/Man-hour functions. Of course, if input/output data could be collected more often, these extrapolative errors could be reduced.

In general, it is no doubt true that the above described indices cannot be as accurate as monthly physical output indexes. On the other hand, the inclusion of the KWH input ought to make possible more accurate interpolations and extrapolations of output/labor relations, either within or beyond established annual historical levels. Finally, the basic technique applied to various sectors of the economy, can be useful in areas where only employment data are currently available. Obviously, additional and different inputs might be required in specific situations, to be combined in a way not yet conceived.

EFFECT OF VARYING DEGREES OF TRANSITORY INCOME ON INCOME ELASTICITY OF EXPENDITURES*

MARILYN DUNSING
University of California, Davis
AND
MARGARET G. REID
University of Chicago

Expenditure-income relations of selected groups of families are examined in two ways: (a) with the importance of transitory income increased and (b) with its importance decreased. The importance of the transitory income is increased by dividing groups of families into subgroups differing in permanent income, and then examining consumption in relation to current income of the subgroups. Such stratification tends to increase the importance of transitory income, since the more homogeneous the families are in permanent income the more the difference in income observed is transitory. The income elasticity of total expenditures is generally lower for the subgroups than for the entire groups—by 30 to 40 per cent. Similar tendencies also occur for expenditures for food, clothing and other consumer goods.

The effect of the contrary tendency, i.e., reduced importance of transitory income, is tested by using mean income of consecutive years as a measure of economic status in place of current annual income. When this is done the coefficient of elasticity of expenditures in relation to income among families is increased over that observed for expenditures in relation to annual income. The increase is around 30 per cent.

The findings support the hypothesis that much of the difference observed among groups in expenditure-income relations is the result of variation in the mixture of the income components.

FAMILY income in any one year is looked upon as having two components, described here as permanent and transitory. The permanent component can suitably be looked upon as the systematic component that families count upon to finance their consumption level.[1] It seems highly probable that expenditures in any one year are more closely related to the permanent than to the transitory component. If so the more important the transitory income the less will be the relationship between expenditures and income, and the lower will tend to be the elasticity of expenditures in relation to income.

This paper presents findings on the relation of annual expenditures of two groups of households to income, under conditions that vary the importance of transitory income, the variation being achieved in the manner described above. The findings support the assumption concerning the relation of income components to consumption.

---

* This article presents a refinement and elaboration of the analysis by Marilyn Dunsing, "Income-Expenditure Relations of Farm Families Using Different Stratification," unpublished doctoral dissertation, Library University of Chicago, Aug. 1954. Assistance from the California Agricultural Station made possible the refinement of the analysis.

[1] Milton Friedman, *A Theory of the Consumption Function*, National Bureau of Economic Research, Princeton University Press, Princeton, 1957, presents a general interpretation of these concepts.

1. PREVIOUS ANALYSES OF ENTIRE GROUPS AND RESPECTIVE SUBGROUPS WITH COMMON INCOME EXPERIENCE

W. W. Cochrane[2] appears to have been the first investigator to have explored the effect of expenditure-income relations of stratification into subgroups with a common income experience. He compared, however, the income elasticity of expenditures of an entire group in 1940 with that of its sub-groups in 1942 that had had a common income in 1940; in other words, his comparison was between the entire group in one year and sub-groups in a later year. Cochrane referred to the expenditure-income relation of the entire group in 1940 as static and that of the sub-groups in 1942 as dynamic.The income elasticities of expenditures of the sub-groups in 1942 were below the elasticity of the entire group in 1940. Cochrane interpreted his findings as follows:

> . . . It is . . . asserted, given a change in income, that families do not immeditely reorganize their budget to conform to the norm of the new income position; hence, expenditures . . . in a dynamic situation are likely to be smaller, income class by income class, as we ascend the income scale, than in a static situation. These continuous data present tangible evidence of this tendency.[3]
>
> . . . And when we find, as we have in these data, that the elasticities of expenditure . . . are extremely inelastic—that is, unresponsive to changes in income—then we must conclude that expenditures for family living are practically constant in the short run. . . .[4]

If the hypothesis to be tested in this paper is valid, what Cochrane observed was not static expenditure-income relations in 1940, in the sense of a long-run adjustment, versus lack of adjustment of expenditures to change in permanent income in 1942. Rather, he observed the effect of some transitory component in the 1940 incomes of the entire group versus the effect of a larger transitory component in the 1942 incomes of the subgroups.

2. DATA USED AND THE EARLIER ANALYSIS

The data used come from records of incomes and expenditures kept by farm families during three consecutive years. They are for the years 1940, 1941, and 1942, and were assembled by Cochrane from state Agricultural Experiment Stations and from the former Farm Security Administration (now the Farmers Home Administration). In this paper families reporting to these two agencies will be referred to as Type I and Type II families, respectively.[5] Their character

[2] W. W. Cochrane, "Farm family budgets—a moving picture," *The Review of Economics and Statistics*, XXIX (Aug. 1947), 189–98.

[3] *Ibid.*, p. 195.

[4] *Ibid.*, p. 197.

[5] Income and expenditure items in the account forms are not entirely comparable for the two groups of families. Since the present investigation relates to sub-groups versus the entire group of which they are a part, the effect of different definitions is not especially relevant. The difference does, however, reduce the usefulness of comparison between Type I and Type II families. It is of interest to note that the purchase of consumer durables tended to be classed as savings by Type II families and as consumption expenditures by Type I families.

Some differences between Type I and Type II families may influence the expenditure-income relations. Type I families had incomes above the average of their areas. They included both owner and tenant operators, and there was some tendency for an over-representation of older, well-established operators. Their accounts were, however, wholly voluntary, being submitted with the thought that efficiency of farm operation should be increased by the advice received from farm management specialists who had examined the accounts. On the other hand, Type II families had incomes about average of their areas. All were relatively young owner-operators who had only recently acquired ownership of their farms by means of a Federal loan. They were required to prepare a budget or spending plan covering both the family and the farm, and accounts were in some measure under scrutiny. These unusual circumstances of Type II families make us suspect that their expenditure-income relations probably had less in common with those of farm families in general than do the expenditure-income relations of Type I families. Consequently, in this paper the data of Type I families are examined more intensively than those of Type II families.

istics are described in some detail in *The Changing Composition of Family Budgets for Selected Groups of Corn Belt Farmers, 1940–42* by Cochrane and Grigg.[6] This analysis uses the following data: those for 236 Type I families in Illinois and Minnesota; and those for 587 Type II families in Illinois, Iowa, and Minnesota. All these families reported a positive income in each of the three years.[7]

Several analyses of these data have already been made. One by Cochrane has been commented on in Section 1 above.[8] An analysis by Reid showed that year-to-year variability of farm operating expenses tends to affect the expenditure-income relationship observed, and that these account-keeping farm families in this respect are very similar to other farm families. Farm operating expenses may cause net income to be atypically low for some families and atypically high for others. A low income caused by relatively high farm operating expenses tends to lower the income elasticity and a high income caused by relatively low farm operating expenses tends to increase it.[9]

Reid has also examined the correlation of incomes of various groups of families in consecutive years and found marked year-to-year variability in relative income positions. Furthermore, the lower the correlation of incomes in consecutive years, the lower tended to be the income elasticity of expenditures. It seemed reasonable to assume that the lower the correlation of incomes, the greater the importance of the transitory components of income, and the greater the importance of transitory incomes, the lower the income elasticity of expenditures.[10]

In an earlier article analyzing these data Reid and Dunsing concluded that variation in level of expenditures among groups of families differentiated by degree of income change is probably caused by the presence of transitory income, for some groups positive and for others negative, rather than by lag in expenditures behind change in permanent income.[11] This paper extends the analysis of these data by testing the effect on income elasticities of varying mixtures of the permanent and transitory income components.

3. ESTIMATES MADE FOR ENTIRE GROUPS AND RESPECTIVE SUB-GROUPS

For Type I families income elasticities of expenditures for 1941 and 1942 are estimated for both entire groups and for their respective sub-groups. The sub-

[6] Bureau of Agricultural Economics, United States Department of Agriculture, 1946. This report and *The Review of Economics and Statistics*, XXIX, Aug., 1947, *loc. cit.* describe the income and expenditure levels of the families included in this analysis. Most of the estimates presented in this paper relate to ungrouped data.

[7] The original set consisted of 367 Type I families and 642 Type II families. Sixteen families were excluded because a negative income was reported for at least one of the three years. The main exclusion was Type I families reporting from Iowa. They were excluded because of extremely high year-to-year variability of incomes. The variability was so high that incomes of any one year seemed likely to have little or no correlation with permanent incomes. Such correlation was a necessary condition in order to ensure subgroups with a transitory income of greater importance than that of the entire group. Type II families of Wisconsin were excluded because of the small number of families.

[8] The tabulation and analysis of the data by Cochrane go far beyond the comparison of income elasticities of expenditures of entire groups and respective sub-groups described in Section 2.

[9] "Effect of Income Concept upon Expenditure Curves of Farm Families," *Conference of Research in Income and Wealth*, National Bureau of Economic Research (Cambridge Press, 1952), XV, pp. 131–74.

[10] For a preliminary report see Margaret G. Reid, "Savings by Family Units in Consecutive Periods," W. W. Heller, ed.), *Savings in the Modern Economy, Symposium*, Minneapolis, 1953, pp. 218–20. Further details are reported by Friedman, *op. cit.*, p. 185, 191–3.

[11] "Effect of Variability of Incomes on Level of Income-Expenditures Curves of Farm Families," *Review of Economics and Statistics*, XXXVIII, Feb. 1956, pp. 90–5.

groups for 1941 are those with 1940 incomes falling within specified intervals. (See Table 352, Panel A.) For 1942 there are two sets of sub-groups: (a) those with common income interval in 1940 and (b) those with common income interval in 1941. (See Table 352, Panel B.) For each of the three sets there are at least seven sub-groups, most of which have an income range of $500 in the common income interval year; consequently, the number of families in any one sub-group is small.[12]

For Type II families income elasticities of expenditures for 1942 only are estimated. There are two sets of sub-groups for 1942: (a) those with common income interval in 1940 and (b) those with common income interval in 1941. The intervals have an income range of $1,000. With such income intervals there are only three sub-groups in each set.[13] (See Table 353.)

4. FORM OF THE EXPENDITURE-INCOME RELATION

Preliminary examination of the data indicated that on logarithmic scales total expenditures in relation to income are fairly linear. The equation is $\log Y = \log a + b \log X$ where X represents annual net money income and Y represents annual cash expenditures for (a) all consumer goods, (b) food, (c) clothing, (d) household, (e) medical care, and (f) other items. For Type I families all coefficients are estimated using ungrouped data and for Type II families they are estimated using grouped data.[14]

5. INCOME ELASTICITIES OF TOTAL EXPENDITURES OF ENTIRE GROUPS AND RESPECTIVE SUB-GROUPS

Type I Families:[15] For Type I families the income elasticities of total expenditures of entire groups was uniformly higher than that of the subgroups. They were as follows:

Year of expenditure-income relation	Entire group	Sub-groups	
	Income elasticity of expenditures	Mean income elasticity of expenditure	Year of common-income interval
1941	.394	.242	1940
1942	.472	.337	1940
1942	.472	.326	1941

In addition, the income elasticities of total expenditures for 1941 and 1942 for the majority of the sub-groups of each set are lower than for the entire group.[16]

[12] For each set the selected sub-groups include over 80 per cent of the 236 families.

[13] The selected sub-groups with common income interval in 1940 include 94 per cent of the 587 families and those with common income interval in 1941, 89 per cent.

[14] Data are grouped by $1,000 class intervals.

[15] All findings presented for Type I families are either taken directly from, or derived from, Table 352.

[16] For 1941 there are seven sub-groups; in six of these the income elasticity of total expenditures is less than that of the entire group. For the two sets of sub-groups for 1942 the results are similar. Thus the relationship between the elasticities of the sub-groups and that of the entire group is consistent even though the number of families in the various sub-groups is small. Such exceptions as occur appear to be random.

TABLE 352

AVERAGE INCOME ELASTICITY OF EXPENDITURES[a] OF VARIOUS CONSUMER GOODS, 1941 AND 1942, OF TYPE I FARM FAMILIES, ENTIRE GROUPS AND RESPECTIVE SUB-GROUPS

Groups and sub-groups	Number of families	Category of consumer goods					
		Total	Food	Clothing	House-hold	Medical care	Other
Panel A. Type I Farm Families, 1941							
Entire group in 1941	236	.394	.185	.375	.387	.470	.442
1. Sub-groups with 1940 income[b]							
a. $1,000 to under $1,500	21	.187	.098	.412	.261	.371	.040
b. $1,500 to under $2,000	32	.223	.080	.340	.180	.059	.169
c. $2,000 to under $2,500	47	.161	.184	.173	.406	.259	.007
d. $2,500 to under $3,000	26	.463	.348	.547	.336	1.004	.560
e. $3,000 to under $3,500	24	.151	.117	.159	.262	−.305	.173
f. $3,500 to under $4,000	27	.302	−.061	−.180	.311	.516	.616
g. $4,000 to under $5 000	22	.260	.369	.431	−.041	.810	.296
Weighted mean of sub-groups	199	.242	.150	.253	.266	.364	.243
Panel B. Type I Farm Families 1942							
Entire group in 1942	236	.472	.205	.498	.290	.604	.636
1. Sub-groups with 1940 income of							
a. $1,000 to under $1,500	21	.292	.161	.318	.236	1.007	.338
b. $1,500 to under $2,000	32	.283	.169	.393	.233	.885	.076
c. $2,000 to under $2,500	47	.406	.037	.301	.189	−.387	.797
d. $2,500 to under $3,000	26	.292	.242	.343	−.143	.140	.431
e. $3,000 to under $3,500	24	.333	.101	−.034	.327	.251	.668
f. $3,500 to under $4,000	27	.471	.206	.590	−.186	.944	.574
g. $4,000 to under $5,000	22	.206	.100	.396	.092	.464	.207
Weighted mean of sub-groups	199	.337	.136	.332	.113	.385	.474
2. Sub-groups with 1941 income of							
a. $1,000 to under $2,000	15	.359	.169	.526	.209	.449	.449
b. $2,000 to under $2,500	21	.072	−.039	.086	.184	.044	.081
c. $2,500 to under $3,000	31	.244	.105	.449	.178	.596	.324
d. $3,000 to under $3,500	28	.396	.166	.228	.657	−.035	.462
e. $3,500 to under $4,000	19	.281	.018	.036	.118	.028	.569
f. $4,000 to under $4,500	21	.353	.025	.438	.179	−.198	.575
g. $4,500 to under $5,000	22	.361	.113	.408	−.179	.503	.528
h. $5,000 to under $6,000	33	.474	.353	.535	−.058	1.185	.855
Weighted Mean of sub-groups	100	.326	.130	.350	.163	.377	.495

[a] All estimates of income elasticities relate to money expenditures in relation to money income.

[b] The income used in classification includes the value of home production and of housing received without direct expense; this is the same procedure as that used by Cochrane. The same relationship would, however, be obtained between entire groups and sub-groups if families had been divided on the basis of net money income.

Source: See text, Section 3.

Regressions of expenditures on income for the entire group and a specified subgroup for 1941 and 1942 are illustrated in Figure 355. For 1941 curve AB representing the entire group has an elasticity of .394, and curve CD representing the sub-group with common income interval of $1,500 to under $2,000 in 1940 has an elasticity of .223. For 1942 curve EF representing the entire group

TABLE 353

AVERAGE INCOME ELASTICITY OF EXPENDITURES[a] OF VARIOUS CONSUMER GOODS, 1942, OF TYPE II FARM FAMILIES, ENTIRE GROUP AND SUB-GROUPS

Groups and sub-groups	Number of families	Category of consumer goods				
		Total	Food	Clothing	Household	Other
Entire group in 1942	587	.383	.252	.306	.363	.627
1. Sub-groups with 1940 income[b] of						
a. $1,000 to under $2,000	259	.335	.124	.267	.304	.611
b. $2,000 to under $3,000	225	.278	.200	.222	.291	.385
c. $3,000 to under $4,000	69	.277	.209	.147	.175	.616
Weighted mean of sub-groups	553	.305	.166	.234	.283	.520
2. Sub-groups with 1941 income of						
a. $1,000 to under $2,000	120	.332	.140	.367	.427	.762
b. $2,000 to under $3,000	260	.263	.170	.182	.153	.425
c. $3,000 to under $4,000	145	.337	.211	.129	.264	.591
Weighted mean of sub-groups	525	.299	.174	.210	.246	.548

[a] See Table 352 for an explanation.
[b] See Table 352 for an explanation
Source: See text, Section 3

has an elasticity of .472 and curve GH representing the sub-group with common income interval of $3,000 to under $3,500 in 1940 has an elasticity of .333 (Panel B).[17]

Subsidiary findings of special interest which relate to total expenditures include: (1) The income elasticity of total expenditures is lower for the entire group in 1941 than in 1942; the same is true of the mean elasticities of the sub-groups. (2) The mean elasticities of the two sets of sub-groups in 1942 are essentially the same, namely, .337 for the sub-groups with common income interval in 1940 and .326 for those with common income interval in 1941.

(3) The mean elasticity of the sub-groups appears to be a function of the absolute level of the elasticity of the entire group. It seems reasonable to assume that the greater the importance of the transitory component in the incomes of the entire group, the greater its importance in the incomes of the respective sub-groups. The difference between the mean income elasticities of the sub-groups and the elasticity of the entire group is essentially the same for the three sets; the method of classification appears to have added about the same amount of transitory component to the incomes. The differences are as follows:

[17] The sub-groups selected have the following characteristics: (1) They have income elasticities which are representative of the mean elasticities of the sub-groups of their respective sets. (2) They illustrate one and two year intervals between the common income interval year and the year of income elasticity. (3) They illustrate positions nearer the lower and upper end of the income size distribution for their respective entire group of families in the common income interval year.

	Difference between the mean elasticity of sub-groups and the elasticity of the entire group
Sub-groups of 1941	−.152
Sub-groups of 1942 with common income interval	
a. in 1940	−.135
b. in 1941	−.146

(4) The ratio of the mean income elasticities of the sub-groups to the elasticity of the entire group is somewhat lower for the 1941 set than for the two 1942 sets. The ratios are as follows:

	Ratio of the mean elasticity of sub-groups to the elasticity of the entire group
Sub-groups of 1941	.614
Sub-groups of 1942 with common income interval	
a. in 1940	.714
b. in 1941	.691

Type II Families:[18] Tendencies for Type II families were similar to those for Type I families. The income elasticity of expenditures of the entire group of Type II families for 1942 was higher than the mean elasticity for the subgroups. It was .383 for the entire group and .305 for the sub-groups with common income interval in 1940 and .299 for the sub-groups with common income interval in 1941. In addition the income elasticity was higher for the entire group than for any of the sub-groups.

Subsidiary findings for Type II families relating to total expenditures include: (1) the mean elasticities of the two 1942 sets of subgroups are very similar, i.e., those with common income intervals in 1940 and those with common income intervals in 1941.

(2) The difference between the mean income elasticities of the sub-groups and the elasticity of the entire group is −.078 for the sub-groups with common income interval in 1940, and −.084 for those with common income interval in 1941. (3) The ratio of the mean income elasticities of the sub-groups to the elasticity of the entire group is .796 for the sub-groups with common income interval in 1940, and .781 for those with common income interval in 1941. These findings of smaller differences and higher ratios for Type II than for Type I families for 1942 are to be expected since the intervals for Type II families have a larger income range than those for Type I families.[19]

[18] All findings presented for Type II families are either taken directly from, or derived from, Table 353.

[19] The larger the income range in the common income interval year, the less the tendency for the families of each of the sub-groups to be alike with respect to the permanent components of income. Hence, in the later year, the less the dispersion of incomes consists of transitory components the less will be the difference in the relative importance of the transitory components of income between the sub-groups and the entire group.

6. INCOME ELASTICITIES OF EXPENDITURES OF VARIOUS CATEGORIES OF CONSUMER GOODS

Type I Families:[20] For five categories of consumer goods the mean income elasticities of expenditures of the sub-groups of Type I families for 1941 and

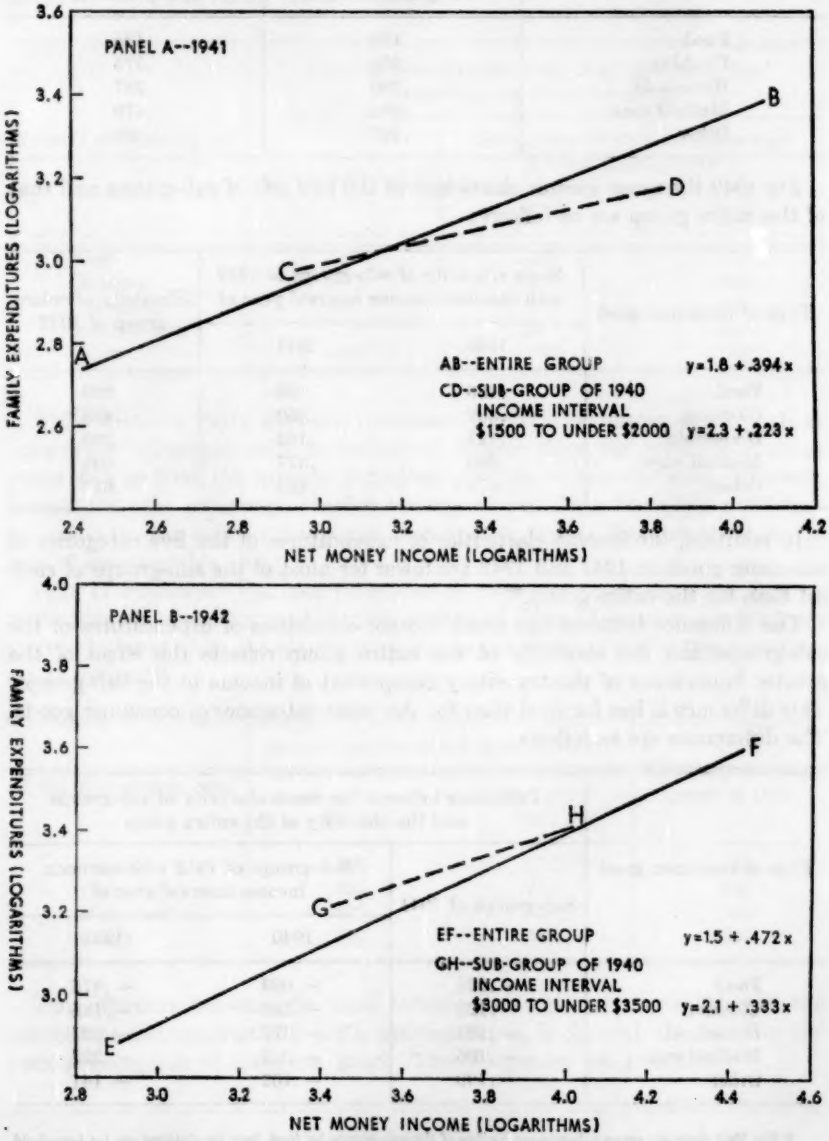


FIG. 355. Family Expenditures in relation to Net Money Income, Type I families. Entire group and a selected subgroup, 1941 and 1942.

[20] All findings presented for Type I families are either taken directly from, or derived from, Table 352.

1942 are uniformly lower than the elasticity of the entire group. For 1941 they are as follows:

Type of consumer good	Mean elasticity of sub-groups of 1941	Elasticity of entire group of 1941
Food	.159	.185
Clothing	.253	.375
Household	.266	.387
Medical care	.364	.470
Other	.243	.442

For 1942 the mean income elasticities of the two sets of sub-groups and that of the entire group are as follows:

Type of consumer good	Mean elasticity of sub-groups of 1942 with common income interval year of		Elasticity of entire group of 1942
	1940	1941	
Food	.136	.130	.205
Clothing	.332	.350	.498
Household	.113	.163	.290
Medical care	.385	.377	.604
Other	.474	.495	.636

In addition, the income elasticities of expenditures of the five categories of consumer goods in 1941 and 1942 are lower for most of the sub-groups of each set than for the entire group.[21]

The difference between the mean income elasticities of expenditures of the sub-groups and the elasticity of the entire group reflects the effect of the greater importance of the transitory component of income in the sub-groups. This difference is less for food than for the other categories of consumer goods. The differences are as follows:

Type of consumer good	Difference between the mean elasticity of sub-groups and the elasticity of the entire group		
	Sub-groups of 1941	Sub-groups of 1942 with common income interval year of	
		1940	1941
Food	−.026	−.069	−.075
Clothing	−.122	−.166	−.148
Household	−.121	−.177	−.127
Medical care	−.106	−.219	−.227
Other	−.199	−.162	−.141

[21] For 1941 there are seven sub-groups; for five of the sub-groups for food, four for clothing, six for household, four for medical care, and five for other the income elasticities of expenditures of the various categories of consumer goods are less than the elasticity of the entire group. For the two sets of sub-groups for 1942 the results are similar. As in the case of total expenditures the exceptions to the tendency of the income elasticities of sub-groups to be less than that of the entire group, for the various categories of consumer goods, appear to be random.

When the difference between the mean income elasticities of the sub-groups and the elasticity of the entire group is expressed as a ratio, the additional transitory component of income of the three sets of sub-groups appears to affect the income elasticities of the five categories of consumer goods in much the same way. The ratios are as follows:

Type of consumer good	Ratio of the mean elasticity of sub-groups to the elasticity of the entire group		
	Sub-groups of 1941	Sub-groups of 1942 with common income interval year of	
		1940	1941
Food	.859	.663	.634
Clothing	.675	.667	.703
Household	.687	.390	.562
Medical care	.774	.637	.624
Other	.550	.745	.778

While it seems likely that the purchase of certain consumer goods such as durables is influenced more by transitory income than the purchase of other goods such as food, the subsidiary findings provide no basis for such differentiation among the categories. Market controls by 1942 may have had a differential effect on the various categories of consumer goods; there seems to be no way of isolating this effect.

Type II Families:[22] For four categories of consumer goods the mean income elasticities of expenditures of the sub-groups of Type II families for 1942 are uniformly lower than the elasticity of the entire group. For the two sets they are as follows:[23]

Type of consumer good	Mean elasticity of sub-groups of 1942 with common income interval year of		Elasticity of entire group of 1942
	1940	1941	
Food	.166	.174	.252
Clothing	.234	.210	.306
Household	.283	.246	.363
Other	.520	.548	.627

The difference between the mean income elasticities of expenditures of the sub-groups and the elasticity of the entire group is, in general, the same for the various categories of consumer goods. The differences are as follows:

[22] All findings presented for Type II families are either taken directly from, or derived from, Table 353.

[23] In addition, the income elasticities of expenditures of the four categories of consumer goods in 1942 are lower for all of the sub-groups with common income interval in 1940 and for two of the three sub-groups with common income interval in 1941.

Type of consumer good	Difference between the mean elasticity of sub-groups and the elasticity of the entire group, for sub-groups with common income interval year of	
	1940	1941
Food	−.086	−.078
Clothing	−.072	−.096
Household	−.080	−.117
Other	−.107	−.079

As with Type I families, the ratio of the mean income elasticities of expenditures of the sub-groups to the elasticity of the entire group for Type II families indicates that the additional transitory component of income of the sub-groups affects the income elasticities of the four categories of consumer goods in much the same way. The ratios are as follows:

Type of consumer good	Ratio of the mean elasticity of sub-groups to the elasticity of the entire group, for sub-groups with common income interval year of	
	1940	1941
Food	.659	.690
Clothing	.765	.686
Household	.780	.678
Other	.829	.874

7. ELASTICITY OF EXPENDITURES IN RELATION TO CURRENT INCOME AND TO INCOME OF THREE YEARS

If income of three years is a better index than current income of the permanent income of families, and if expenditures are more closely related to permanent than to current income, then one would expect the elasticity of expenditures in any year to be higher in relation to income of three years than to current income. Such a difference is observed for Type I farm families for all categories of consumption. The coefficients of elasticity of expenditures in 1942 in relation to income are as follows:

Expenditures	Current Income	Income of Three Years
Total	.394	.618
Food	.185	.253
Clothing	.375	.588
Household	.387	.463
Medical Care	.470	.767
Other	.442	.818

A similar estimate for total expenditures in 1942 was made for Type II families. It shows the following coefficients of elasticity of 1942 expenditures in relation to income: income of 1942, .383 and three-year income, .508.

Thus for Type I families income elasticity of total expenditures in relation to three-year income was higher by 31 per cent than that in relation to current

income, and of Type II families 32 per cent. For Type I families the tendency for each consumption category was similar to that of total expenditures.

This comparison demonstrates that elasticity of expenditures in relation to annual income, at least for farm families, is likely to understate the difference in expenditures among families that occurs because of difference in level of income. It does not, however, provide an estimate of the elasticity of expenditures in relation to permanent income, since it seems highly probable that transitory income influenced the variability of the three-year incomes.

Furthermore for Type I families the elasticity of expenditures for consumption in 1942 in relation to mean income of three years is about 84 per cent higher than the mean elasticity for the sub-groups stratified by the income of an earlier year, and for Type II families about 70 per cent higher.

8. SUMMARY AND CONCLUDING REMARKS

The general hypothesis of this paper is that the more the difference in incomes of families consists of transitory components of income, the lower will tend to be the elasticity of expenditures in relation to that income. The hypothesis was tested by dividing a group of families into sub-groups that have a common income experience in an earlier year. This method of classifying families causes the transitory components of income in a later year to be more important for sub-groups than for the entire group. It was further tested by estimating expenditures in relation to three-year incomes.

Income elasticities of expenditures observed support the hypothesis. The income elasticity of total expenditures for entire groups is uniformly higher than the mean elasticities of the respective sub-groups. In addition, for various categories of consumer goods the income elasticity of expenditures for entire groups is also uniformly higher than the mean elasticities for the respective sub-groups. Furthermore income elasticities of expenditures in relation to three-year incomes was higher than those in relation to one-year incomes.

Higher elasticities for entire groups than for respective sub-groups indicate that the more alike families are in the permanent components of their income, the more the difference in income reported in any one year is likely to be caused by transitory income. Thus the income elasticity of expenditures of a single occupational group with a relatively narrow range of permanent incomes is likely to be lower than the income elasticity of expenditures of all occupational groups combined unless the group with the lower dispersion of permanent incomes also has a lower dispersion of transitory incomes.

Probably the most important finding of this analysis is the further evidence it provides that much of the variation in the income elasticity of expenditures observed among groups of families is a function of the mixture of permanent and transitory income components. Because of such mixture, the elasticities of expenditures estimated from the data give a biased estimate of the relation of expenditures to the permanent component of income and of the effect on expenditures of a recent change in the permanent component of income. While the elasticities presented in this paper are estimated from data of farm families, similar results would be expected for other groups of families whose incomes also contain considerable transitory components. It seems highly probable that all of the income elasticities of expenditures reported in this analysis understate the income elasticity of expenditures in relation to the permanent component of income.

THE FIRST 1,945 BRITISH STEAMSHIPS

J. R. T. Hughes and Stanley Reiter
Purdue University

This paper presents data showing the main characteristics of all the steamships built in the United Kingdom and placed under permanent British registry from 1814 through 1860. The historical setting in which the British steam merchant fleet developed is discussed and the data are interpreted in relation to that environment. It is argued on the basis of these data that many of the prevailing notions about steam shipping before 1860 need to be revised.

1. INTRODUCTION

This paper presents data, which have not been utilized before, showing the main characteristics of all the steamships built in the United Kingdom and placed under permanent British registry[1] from 1814 through 1860. These data should provide scholars the means to gain a clearer perspective of the history of steam navigation prior to the outbreak of the American Civil War.

The only official data on British steamship construction which have been generally available to scholars heretofore have been the "Number and Tonnage" figures given in the *Statistical Abstract for the United Kingdom* from 1840 onwards. These *Statistical Abstract* data are aggregates enumerating all steamships built and first registered in the U.K. and contain no information regarding how much of annual construction actually passed into permanent U.K registry. Nor do these series contain any information regarding the number of wood and iron steamers built, or the types, whether paddle or screw propelled. The new series presented in this paper fill all of these gaps.[2] In addition, measures have been constructed which show carrying capacity of new construction entered into permanent U.K. registry. These measures take into account advances in the technology of shipbuilding.

The present paper presents information which indicates that many of the prevailing notions about steam shipping before 1860 need to be rather sharply revised. It is shown, for example, that the general adoption of iron in steamship construction dates from the 1840's, more than a full decade earlier than the accepted estimates; that by the end of the 1850's most of the British steam merchant fleet was iron rather than wood; that the iron-screw steamer was

[1] The data presented below are not the same as annual data given in the *Statistical Abstract for the United Kingdom* for ships "built and first registered in the United Kingdom." Such data include ships which, although first built in the U.K., later passed out of British registry by reason of sales to foreigners. As is indicated below, the ships included in the data presented in this paper were entered permanently into U.K. registry during the period covered by these data. Hence in some years sales of British-built steamers to foreigners is reflected in the difference between the first registrations and the permanent registrations indicated by the data below. Similarly, because of British purchases of steamers from abroad (which might originally have been British built but which had temporarily passed out of British registry), additions to the British steam merchant fleet may in some years seem to be slightly larger than was warranted by current U.K. construction. For details of British ship registry practices see: Ernest W. Blocksidge, *Hints on the Register Tonnage of Merchant Ships*, Liverpool, 1942, pp. 9–13.

[2] The authors are indebted to the Cunard Steamship Company Ltd., for providing data on early Cunard Steamers which could be used in checking the accuracy of the transport capacity measures in Table 367 below, and for information on the obsolescence of early steamers.

predominant in new British steamship construction from at least 1851 onwards; that the 1853 trade boom together with the stimulus of the Crimean War demand for shipping were decisive in introducing the iron-screw steamer; that the iron-screw steamer very largely supplanted both wood and iron paddle steamers by the end of the Crimean War. On the basis of this evidence including our measure of transport capacity it will be argued that contemporaries were probably correct in their estimates of the extent of the steamship's penetration of the shorter ocean-freight routes in the 1850's and early 1860's—estimates which have been thought to be open to serious doubt. It is argued that Britain's new steam merchant marine was a powerful influence in the development of Britain's exceedingly strong balance of payments position in the third quarter of the 19th Century.

2. THE DATA

At the order of the 1861 session of Parliament a return was printed containing certain information covering all of the "Steam Vessels" permanently registered in the merchant fleet of the U.K. on or before January 1, 1861.[3] The following information is contained in this Parliamentary Paper: the length, breadth, and depth of hold of each ship is given, together with its gross and net tonnages, and the indicated horse power of its engines. The type of each ship is given, whether it was wood paddle, wood screw, iron paddle, iron screw or steel. Finally, the date of build, port of registry, and name of the registered builder (or builders) for each ship is stated.

There are 1,945 steam vessels listed in this fashion as "the whole" number of British steamers permanently registered in the U.K. from 1814 through 1860, with a total of 440,999 net and 692,101 gross tons.[4] The data are continuous from 1823 onwards and hence the analysis below concentrates on the period 1823–1860.

It is not difficult to understand why these data have not been utilized before, considering their organization—or rather, their lack of it. The ships were given numbers by 95 British ports of registry from London to Wexford. This

[3] Full title: "A RETURN, 'in a Tabular Form, with Consecutive Numbers, of the whole of the Steam Vessels Registered in the United Kingdom on or before the 1st day of January 1861; stating in separate Columns, the following Particulars:—Official Number of Vessel; Vessel's name; Port of Registry; Date of Registry; Date of Build; Registered Owners; Dimensions of Vessels in Length and Breadth, and Depth of Hold; Tonnage (exclusive of Engine Room); and Gross Tonnage, distinguishing Vessels built of Iron and Vessels having Screw Propellers; also distinguishing Vessels measured under the Merchant Shipping Act of 1854 from those measured under previous Acts; and Estimated Horse Power of their Engines, and giving the Aggregate Number of Vessels and Amount of Tonnage; with an Index for easy reference attached to it, giving the Names of the Vessels in Alphabetical Order, with Numbers to each, corresponding with the Consecutive Numbers in the Return (in continuation of Parliamentary Paper No. 449, of Session 1860)'," *Accounts and Papers*, 1861, LVIII (371).

[4] The "net tonnage" figures in Tables 363 and 367 below are the official "gross tonnage" measurements less tonnage assigned to the engine rooms of the ships. Gross tonnage is generally "the measurement of the under-deck tonnage, with the addition of the 'tween-deck spaces situated above the tonnage deck and all 'enclosed' spaces above the upper deck." Blocksidge, *op. cit.*, p. 7. Of the 1,944 ships included in Tables 363 and 367, 94 ships lacked entries in one or more of the categories "net tonnage," "gross tonnage," "indicated horsepower." Estimates of the missing measurements were based on corresponding measurements for other ships built in the same year, or as nearly in the same year as possible (and if possible, ships having the same owners), and having the same dimensions and known tonnages and horsepower as the ships in question. It was felt that this procedure yielded more accurate estimates than would the standard procedure of using linear regressions. The error introduced by these estimates is no doubt an insignificant part of the total. Tables 363 and 367 include river steamers; these are not included in certain *Statistical Abstract* data on steamships. Because these ships were all officially registered, they were of greater than 15 net register tons; ships of less than that tonnage (or 15 tons burden under the older British Registry legislation) need not have been officially registered. Blocksidge, *op. cit.*, pp. 9–10.

serial ordering begins with London, presumably because it had the greatest number of registries, and then proceeds alphabetically from Aberystwith to Yarmouth for 63 English and Welsh ports. The numbers continue from Aberdeen to Wigtown for Scottish ports and from Ballina to Wexford for Irish ports. A special index follows with the names of all the ships in alphabetical order. Each ship entry includes the official British registry number, but one can conveniently find an individual ship only if one knows its name; the registry numbers are not presented serially. No attempt was made to order the ships other than by port of registry. To use these data for different purposes involved their complete reorganization at each stage of analysis; 1,945 individual ships with at least 13 entries per ship—a considerable task if done by hand. Punch-card calculators were used by the present authors for the thousands of computations involved in organizing these data.[5]

3. GROWTH OF THE BRITISH STEAM MERCHANT FLEET 1823–1860

In Table 363 the data are organized in continuous annual series running from 1823[6] through 1860. By the end of 1860 a total of 1,945 steam vessels had been entered into permanent British registry with 692,101 gross tons.[7] It is seen that the growth of the steam merchant fleet was slow until the early 1840's, a peak was reached in 1841, another in 1847 and in 1850 there began a strong upward surge in additions to the fleet which reached its peak in 1855, the final year of the Crimean-War boom in shipbuilding. A substantial recovery became apparent only in 1860.

These movements in additions to the steam merchant fleet roughly correspond to changes in the data for total ships (both sail and steam) built and first registered in the United Kingdom. This may be seen in Chart 365 where data for the volume of foreign trade are compared to data for ships built and first registered in the U.K. The upward trend in foreign trade (and, *ceteris paribus*, in total freight revenues) no doubt accounts in large part for the accompanying upward trend in the tonnage of ships built and first registered in Britain. The rise in the trade data accelerates appreciably in the late 1840's and in the 1850's and is accompanied by a sharp upturn in shipbuilding. We will turn to this in detail below, since it was in this trade upswing that the iron steamer came into its own. The boom in shipbuilding centering upon 1840–1841 is evident in both the data for total ships built and registered and in those for steamers entered into permanent British registry (cf. Chart 365 and Table 363) except that the peak in total ships built and first registered comes in 1840 while the peak for steamers added to the British merchant fleet comes in 1841. Whereas

[5] The authors are indebted to the Computing Laboratory of Purdue University for assistance with the computations. We would be glad to provide at cost a set of punched cards containing the original data to any scholar interested in making further use of these data.

[6] According to these returns a single steamer, the "Industry" was entered permanently into British registry in 1814. The port of registry was Glasgow; the owner was A. McGeorge. This ship was 68 feet long, 17 feet wide, 8 feet deep, net tonnage was 43, gross tonnage 69, and indicated horsepower was 14. No other steamers appear until 1823 when 4 were put under permanent registry.

[7] Including the adjustments in the data noted above. In 1860 the *Statistical Abstract* lists 2,000 as being the number of steam vessels "registered as belonging" to the United Kingdom and the register tonnage is given as 454,327. The differences between these figures and those in Table 363 are easily accounted for. The *Statistical Abstract* data include vessels registered in the channel islands, whereas the data in Table 363 include only registry ports on the main islands (i.e., Great Britain and Ireland) and the *Statistical Abstract* data are *register tonnages*. The latter are significantly lower than *gross tonnage* estimates. See Blocksidge, *op. cit.*, pp. 5–9 on tonnage measurements.

TABLE 363

DISTRIBUTION OF STEAMSHIPS PERMANENTLY REGISTERED IN THE UNITED KINGDOM BY TYPE, NUMBER, NET AND GROSS TONNAGE, FOR EACH YEAR FROM 1823 TO 1860

Year of Build	Wood Steamships								
	Paddle			Screw			Total		
	No.	Net Tons	Gross Tons	No.	Net Tons	Gross Tons	No.	Net Tons	Gross Tons
1823	4	427	817	—	—	—	4	427	817
1824	4	341	682	—	—	—	4	341	682
1825	5	419	894	—	—	—	5	419	894
1826	11	964	1,831	—	—	—	11	964	1,831
1827	3	451	662	—	—	—	3	451	662
1828	3	63	141	—	—	—	3	63	141
1829	4	301	613	—	—	—	4	301	613
1830	5	443	752	—	—	—	5	443	752
1831	4	447	725	—	—	—	4	447	725
1832	8	910	1,491	—	—	—	8	910	1,491
1833	8	503	1,222	—	—	—	8	503	1,222
1834	17	2,350	4,042	—	—	—	17	2,350	4,042
1835	20	2,273	4,187	—	—	—	20	2,273	4,187
1836	18	2,468	4,440	—	—	—	18	2,468	4,440
1837	29	3,632	6,285	—	—	—	29	3,632	6,285
1838	31	1,789	3,547	—	—	—	31	1,789	3,547
1839	16	1,473	2,732	—	—	—	16	1,473	2,732
1840	23	2,936	5,572	1	292	355	24	3,228	5,927
1841	27	9,265	14,644	—	—	—	27	9,265	14,644
1842	22	4,734	8,465	—	—	—	22	4,734	8,465
1843	27	1,547	3,233	—	—	—	27	1,547	3,233
1844	23	2,481	4,631	—	—	—	23	2,481	4,631
1845	15	927	1,806	—	—	—	15	927	1,806
1846	26	2,027	3,719	1	163	240	27	2,190	3,959
1847	47	4,396	8,735	—	—	—	47	4,396	8,735
1848	46	2,506	5,399	—	—	—	46	2,506	5,399
1849	22	2,177	3,815	—	—	—	22	2,177	3,815
1850	12	3,525	6,268	2	549	698	14	4,074	6,966
1851	20	5,179	8,758	1	103	151	21	5,282	8,909
1852	35	2,669	5,319	2	180	265	37	2,849	5,584
1853	31	3,571	6,239	3	221	321	34	3,792	6,560
1854	26	2,326	4,206	3	1,527	2,189	29	3,853	6,395
1855	28	3,235	5,328	2	162	185	30	3,397	5,513
1856	43	2,768	5,719	2	163	232	45	2,931	5,951
1857	55	1,066	4,057	5	163	313	60	1,229	4,370
1858	33	2,245	4,710	2	64	112	35	2,309	4,822
1859	39	771	2,619	5	858	1,182	44	1,629	3,801
1860	35	1,096	2,769	5	195	275	40	1,291	3,044

Year of Build	Iron Steamships								
	Paddle			Screw			Total		
	No.	Net Tons	Gross Tons	No.	Net Tons	Gross Tons	No.	Net Tons	Gross Tons
1823	—	—	—	—	—	—	—	—	—
1824	—	—	—	—	—	—	—	—	—
1825	—	—	—	—	—	—	—	—	—
1826	—	—	—	—	—	—	—	—	—
1827	—	—	—	—	—	—	—	—	—
1828	—	—	—	—	—	—	—	—	—
1829	—	—	—	—	—	—	—	—	—
1830	—	—	—	—	—	—	—	—	—
1831	—	—	—	—	—	—	—	—	—
1832	—	—	—	—	—	—	—	—	—
1833	—	—	—	—	—	—	—	—	—
1834	—	—	—	—	—	—	—	—	—
1835	—	—	—	1	368	473	1	368	473
1836	2	194	284	—	—	—	2	194	284
1837	4	390	677	—	—	—	4	390	677
1838	5	174	298	—	—	—	5	174	298
1839	5	295	559	—	—	—	5	295	559
1840	7	576	947	1	1,734	3,509	8	2,310	4,456

TABLE 363—(*continued*)

Year of Build	Iron Steamships								
	Paddle			Screw			Total		
	No.	Net Tons	Gross Tons	No.	Net Tons	Gross Tons	No.	Net Tons	Gross Tons
1841	7	971	1,548	—	—	—	7	971	1,548
1842	10	818	1,319	1	116	171	11	934	1,490
1843	6	669	1,105	—	—	—	6	669	1,105
1844	29	4,025	6,947	3	255	400	32	4,280	7,347
1845	26	3,126	5,483	3	424	651	29	3,550	6,134
1846	17	4,589	7,233	10	2,716	3,803	27	7,305	11,036
1847	31	7,919	12,270	6	1,777	2,458	37	9,696	14,728
1848	19	3,518	5,853	6	1,667	2,999	25	5,185	8,852
1849	20	4,285	6,763	8	2,715	3,919	28	7,000	10,682
1850	18	3,673	5,871	12	2,393	3,547	30	6,066	9,418
1851	18	3,902	5,850	15	6,824	10,180	33	10,726	16,030
1852	15	2,707	4,038	20	12,695	18,984	35	15,402	23,022
1853	20	4,599	7,538	43	21,321	31,455	63	25,920	38,993
1854	25	5,479	9,143	82	39,274	56,310	107	44,753	65,453
1855	35	4,633	8,066	84	42,541	64,261	119	47,174	72,327
1856	32	6,350	10,273	71	27,355	38,567	103	33,705	48,840
1857	42	4,061	8,521	68	27,045	38,432	110	31,106	46,953
1858	34	4,956	8,081	41*	29,263*	41,344*	73*	34,219*	49,425*
1859	30	8,319	13,734	48†	15,546†	21,663†R	78†	23,865†	35,397†
1860	30	6,434	13,462	75R	32,924R	45,461R	105R	39,358R	58,923R

Year of Build	Total Wood and Iron Steamships								
	Paddle			Screw			Total		
	No.	Net Tons	Gross Tons	No.	Net Tons	Gross Tons	No.	Net Tons	Gross Tons
1823	4	427	817	—	—	—	4	427	817
1824	4	341	682	—	—	—	4	341	682
1825	5	419	894	—	—	—	5	419	894
1826	11	964	1,831	—	—	—	11	964	1,831
1827	3	451	662	—	—	—	3	451	662
1828	3	63	141	—	—	—	3	63	141
1829	4	301	613	—	—	—	4	301	613
1830	5	443	752	—	—	—	5	443	752
1831	4	447	725	—	—	—	4	447	725
1832	8	910	1,491	—	—	—	8	910	1,491
1833	8	503	1,222	—	—	—	8	503	1,222
1834	17	2,350	4,042	—	—	—	17	2,350	4,042
1835	20	2,273	4,187	1	368	473	21	2,641	4,660
1836	20	2,667	4,724	—	—	—	20	2,662	4,724
1837	33	4,022	6,962	—	—	—	33	4,022	6,962
1838	36	1,963	3,845	—	—	—	36	1,963	3,845
1839	21	1,768	3,291	—	—	—	21	1,768	3,291
1840	30	3,512	6,519	2	2,026	3,864	32	5,538	10,383
1841	34	10,236	16,192	—	—	—	34	10,236	16,192
1842	32	5,552	9,784	1	116	171	33	5,668	9,955
1843	33	2,216	4,328	—	—	—	33	2,216	4,328
1844	52	6,506	11,578	3	255	400	55	6,761	11,978
1845	41	4,053	7,289	3	424	651	44	4,477	7,940
1846	43	6,616	10,952	11	2,879	4,043	54	9,495	14,995
1847	78	12,315	21,005	6	1,777	2,458	84	14,092	23,463
1848	65	6,024	11,252	6	1,667	2,999	71	7,691	14,251
1849	42	6,462	10,578	8	2,715	3,919	50	9,177	14,497
1850	30	7,198	12,139	14	2,942	4,245	44	10,140	16,384
1851	38	9,081	14,608	16	6,927	10,331	54	16,008	24,939
1852	50	5,376	9,357	22	12,875	19,249	72	18,251	28,606
1853	51	8,170	13,777	46	21,542	31,776	97	29,712	45,553
1854	51	7,805	13,349	85	40,801	58,499	136	48,606	71,848
1855	63	7,868	13,394	86	42,703	64,446	149	50,571	77,840
1856	75	9,118	15,992	73	27,518	38,799	148	36,636	54,791
1857	97	5,127	12,578	73	27,208	38,745	170	32,335	51,323
1858	67	7,201	12,791	43	29,327	41,456	110*	36,528*	54,247*
1859	69	9,090	16,353	53†	16,404†	22,845*	122†	25,494†	39,198†
1860	65	7,530	16,231	80R	33,119R	45,736R	145R	40,649R	61,967R

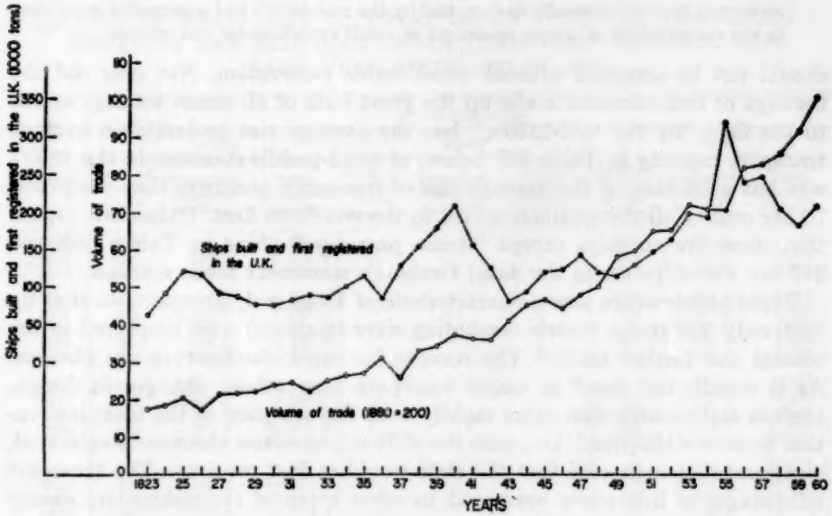
Source: *Accounts and Papers*, 1861, LVIII (371).

* Includes the "Great Eastern," 13,344 net tons and 18,915 gross tons; and 1 steel ship, 79 net tons and 111 gross tons.

† Includes 1 steel ship 129 net tons and 204 gross tons.

R Includes three steel ships, 265 net tons and 398 gross tons.

CHART 365


Ships built and first registered in the U.K. (000 tons)
Volume of trade
Ships built and first registered in the U.K.
Volume of trade (1880 = 200)
YEARS
1823 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59 60

the peaks in ship construction in 1847 and 1855 are easily accounted for (below), the 1840(41) peak in shipbuilding in Britain is not adequately accounted for in the literature and remains largely unexplained. It certainly does not conform to movements in other data of economic activity.[8]

The information contained in Table 363 alone is sufficient to justify a rather sharp revision of accepted interpretations of the growth of the British steam merchant marine up to 1861. It is seen in Table 363 that some 60 per cent of all the gross tonnage of steamships added to the fleet in 1823–60 were added in the seven years 1854–60. Moreover, about 85 per cent of this gross tonnage consisted of iron ships and roughly 70 per cent of iron-screw steamers. As early as 1844 the gross tonnage of iron steamers exceeded that of wood[9] and by 1851 the tonnage of iron-screw steamers alone exceeded total wood steamers (both paddle and screw).[10] Hence the well-known view that "iron had not been generally adopted for steamers before the decade 1855–65"[11] is rather far from the mark. From 1851 onwards, the additions of iron-screw steamers to the fleet led all other types and by 1853 iron-screw steamers clearly dominated all others combined in the growth of the steam merchant marine. Considering the data in Tables 363 and 367 it ought to be noted that G. S. Graham's recent statement

---

[8] See Gayer, A. D., Rostow, W. W., and Schwartz, A. J., *The Growth and Development of the British Economy 1790–1850*, Oxford, 1952, I, pp. 276–303 for other economic data in the period.

[9] 7,347 gross tons of iron steamers compared to 4,631 gross tons of wood. Table 363.

[10] 10,180 gross tons of iron-screw steamers, 5,850 gross tons of iron paddle and 8,909 gross tons of wood steamers of both kinds.

[11] J. H. Clapham, *Economic History of Modern Britain*, Cambridge, 1942, II, p. 63. Clapham's discussion also suggests rather strongly, p. 64, that most of the steam fleet in the 1850's was still wood. L. C. A. Knowles places the beginnings of success for the iron steamship in the transport of passengers and cargo only "Between 1850 and 1860." *The Industrial and Commercial Revolutions in Great Britain during the Nineteenth Century*, London, 1926, p. 298.

". . . iron proved to be a stubborn and uncertain medium, especially in the hands of craftsmen trained to handle timber, and by the mid-1850's had superseded wood only in the construction of steam passenger or small experimental war vessels . . . "[12]

should not be accepted without considerable reservation. Not only did the tonnage or iron steamers make up the great bulk of all steam tonnage added to the fleet "by the mid-1850's," but the average size (reflected in average transport capacity in Table 367 below) of wood-paddle steamers in the 1850's was but a fraction of the average size of iron-screw steamers that comprised 70 per cent of all the steamers added to the merchant fleet. Unless one argues that there are no ships except "steam passenger" ships in Tables 363 and 367 (no war ships are in the data) Graham's statement needs revision.

Rapid obsolescence was a characteristic of these early steamers, so that by 1860 only 929 steam vessels (excluding river steamers) were employed in the coastal and foreign trade.[13] The reasons for rapid obsolescence are obvious. As is usually the case[14] in major transport innovations, changes in design, engines and construction came rapidly once the adoption of the basic innovation became widespread; i.e., once the shift to iron-screw steamers accelerated, and competition insured that the shift would in fact continue. The transport advantages of iron-screw compared to other types of steamships are clearly seen in Table 367 where the average transport capacity[15] of iron-screw steamers was in general much greater by the mid-1850's than that of any other type of steamer.

Of course the rate of obsolescence varied from ship to ship, but it appears that few of the steam ships built before the 1850's remained in service for much longer than a decade at most. The "Brittania" of 1840, the first steam Cunarder in regular trans-Atlantic service, was retired from that service in 1849; and Cunard's famous iron-paddle steamer "Persia," launched in 1856. was retired from the Atlantic service in 1868, after a scant 12 years of service.[16]

Given the high rate of obsolescence and the fact that most of the British steam merchant fleet of 1860 had been built within a decade of that date, it is clear that by 1860 the British steam merchant marine was iron, and moreover, iron screw.[17] The implications of this fact can be appreciated by inspection of the measures of transport capacity in Table 367 below—the earning capacity of this steam fleet was growing prodigiously. In fact, even though by 1860 steamers accounted for only 10 per cent of total British register tonnage, the total freight earnings accruing to these steamers must have been much greater.

---

[12] G. S. Graham, "The Ascendency of the Sailing Ship 1850–85," *Economic History Review* (Second Series), Vol. IX, No. 1, August 1956, pp. 75–6.

[13] *Statistical Abstract.*

[14] Consider the experience of air and highway transportation, as well as the early history of rail transporation.

[15] See next section for a discussion of this measurement.

[16] Private communication to the authors from the Cunard Steamship Company Ltd. It should be emphasized that some steamers in less competitive routes might enjoy a comparatively long service. By the end of the 19th century the rate of obsolescence slowed down significantly as designs and engines had passed through many experimental stages.

[17] There were in fact already 5 steel ships in the registry by 1860 (Table 363). Clapham, *op. cit.*, pp. 56, 62 found his first steel steamships in the merchant fleet only in 1863–65. The data in Table 363 make it possible to avoid repetition of earlier errors concerning the size of Britain's early steam merchant fleet. For example, Gayer, Rostow, and Schwartz, *op. cit.*, p. 286 cite G. R. Porter's *Progress of the Nation* to the effect that by 1842 some 4 per cent of registered British shipping was steam. This is obviously an overstatement. Total tonnage under British registry in 1842 was 2,991,000, 4 per cent of which would be 119,640 register tons; according to the data in Table 363 only about 70,000 *gross* tons of steam shipping had been added to the British merchant fleet by 1842, and much of that must by 1842 have been rendered obsolete and taken out of service.

TABLE 367

RATIO OF TOTAL NET TO TOTAL GROSS TONNAGE, TRANSPORT CAPACITY PER SHIP AND TOTAL TRANSPORT CAPACITY FOR EACH YEAR FROM 1823 TO 1860

Year of Build	Wood Steamships								
	Paddle			Screw			Total		
	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity	Ratio of of Net Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity
1823	.523	990.5	3,990.4	—	—	—	.523	990.5	3,990.4
1824	.500	793.2	3,172.8	—	—	—	.500	793.2	3,172.8
1825	.469	802.6	4,013.2	—	—	—	.469	802.6	4,013.2
1826	.526	795.8	8,754.2	—	—	—	.525	795.8	8,754.2
1827	.681	1,331.0	3,993.0	—	—	—	.681	1,331.0	3,993.0
1828	.447	1,780.0	534.0	—	—	—	.447	1,780.0	534.0
1829	.491	798.5	3,193.9	—	—	—	.491	798.5	3,193.9
1830	.589	781.5	3,907.6	—	—	—	.589	781.5	3,907.6
1831	.616	1,041.9	4,167.7	—	—	—	.616	1,041.9	4,167.7
1832	.610	1,031.2	8,249.7	—	—	—	.610	1,031.2	8,249.7
1833	.412	583.9	4,670.9	—	—	—	.412	583.9	4,670.9
1834	.581	1,374.1	23,359.8	—	—	—	.581	1,374.1	23,359.8
1835	.543	1,008.7	22,192.0	—	—	—	.543	1,008.7	22,192.0
1836	.556	1,298.4	23,371.8	—	—	—	.556	1,298.4	23,371.8
1837	.578	1,252.5	36,321.1	—	—	—	.578	1,252.5	36,321.1
1838	.504	565.8	17,540.0	—	—	—	.504	565.8	17,540.0
1839	.539	923.4	14,774.1	—	—	—	.539	923.4	14,774.1
1840	.527	1,239.0	28,497.8	.822	1,839.6	1,839.6	.545	1,264.1	30,337.4
1841	.633	3,418.8	92,306.9	—	—	—	.633	3,418.8	92,306.9
1842	.559	2,157.5	47,464.7	—	—	—	.559	2,157.5	47,464.7
1843	.480	545.1	14,718.4	—	—	—	.480	545.1	14,718.4
1844	.536	1,124.8	25,871.2	—	—	—	.536	1,124.8	25,871.2
1845	.513	551.2	8,267.9	—	—	—	.513	551.2	8,267.9
1846	.545	732.6	19,047.2	.679	1,173.6	1,173.6	.553	748.9	20,220.8
1847	.503	967.2	45,457.3	—	—	—	.503	967.2	45,457.3
1848	.464	542.5	24,956.6	—	—	—	.464	542.5	24,956.6
1849	.571	1,176.5	25,883.8	—	—	—	.571	1,176.5	25,883.8
1850	.562	3,957.8	47,493.2	.786	2,377.2	4,754.5	.585	3,732.0	52,247.7
1851	.591	3,551.7	71,033.7	.682	916.7	916.7	.593	3,426.2	71,950.4
1852	.502	1,040.7	36,423.1	.679	902.7	1,805.5	.510	1,033.2	38,228.6
1853	.572	1,368.0	42,408.1	.688	719.0	2,156.9	.578	1,310.7	44,565.0
1854	.553	958.0	24,908.7	.698	5,106.0	15,319.7	.602	1,387.2	40,228.4
1855	.607	1,192.9	33,400.4	.876	501.7	1,013.4	.616	1,177.7	34,413.8
1856	.484	783.7	33,697.8	.702	815.9	1,631.7	.493	785.1	35,329.5
1857	.263	205.0	11,275.4	.521	284.8	1,424.1	.281	211.7	12,699.5
1858	.477	855.9	28,243.3	.571	261.6	523.1	.478	821.9	28,766.4
1859	.294	203.8	7,950.1	.726	1,494.7	7,473.5	.428	350.5	15,423.6
1860	.396	300.0	10,499.1	.709	334.9	1,674.5	.424	304.3	12,173.6

Year of Build	Iron Steamships								
	Paddle			Screw			Total		
	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity
1823	—	—	—	—	—	—	—	—	—
1824	—	—	—	—	—	—	—	—	—
1825	—	—	—	—	—	—	—	—	—
1826	—	—	—	—	—	—	—	—	—
1827	—	—	—	—	—	—	—	—	—
1828	—	—	—	—	—	—	—	—	—
1829	—	—	—	—	—	—	—	—	—
1830	—	—	—	—	—	—	—	—	—
1831	—	—	—	—	—	—	—	—	—
1832	—	—	—	—	—	—	—	—	—
1833	—	—	—	—	—	—	—	—	—
1834	—	—	—	—	—	—	—	—	—
1835	—	—	—	.778	3,275.2	3,275.2	.778	3,275.2	3,275.2

TABLE 367—(*continued*)

Year of Build	Iron Steamships								
	Paddle			Screw			Total		
	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity
1836	.683	893.4	1,786.8	—	—	—	.683	893.4	1,786.8
1837	.576	889.5	3,558.1	—	—	—	.576	889.5	3,558.1
1838	.584	287.5	1,437.6	—	—	—	.584	287.5	1,437.6
1839	.528	487.4	2,437.0	—	—	—	.528	487.4	2,437.0
1840	.608	727.4	5,092.1	.494	15,779.4	15,779.4	.518	2,609.0	20,871.5
1841	.627	1,277.3	8,940.9	—	—	—	.627	1,277.3	8,940.9
1842	.620	730.5	7,304.8	.678	626.4	626.4	.627	721.1	7,931.2
1843	.605	992.6	5,955.6	—	—	—	.605	992.6	5,955.6
1844	.579	1,350.3	39,158.3	.637	604.4	1,813.3	.582	1,280.4	40,971.6
1845	.570	1,170.3	30,428.0	.651	1,090.2	3,270.7	.579	1,162.0	33,698.7
1846	.634	2,618.4	44,512.8	.714	2,356.4	23,564.1	.662	2,521.4	68,076.9
1847	.645	2,611.7	80,961.8	.723	2,253.5	13,520.8	.658	2,553.6	94,482.6
1848	.601	1,939.2	36,844.5	.556	2,382.1	14,292.5	.586	2,045.5	51,137.0
1849	.633	2,549.6	50,992.8	.692	3,610.4	28,882.8	.655	2,852.7	79,875.6
1850	.626	2,544.2	45,796.2	.675	2,053.8	24,645.4	.644	2,348.1	70,441.6
1851	.667	2,521.9	45,393.4	.670	4,799.4	71,990.9	.669	3,557.1	117,384.3
1852	.670	2,062.5	30,938.2	.669	6,679.7	133,594.2	.669	4,700.9	164,532.4
1853	.610	2,466.7	49,334.3	.678	5,112.1	219,821.0	.665	4,273.3	269,155.3
1854	.599	2,474.5	61,863.6	.697	4,639.4	380,428.2	.684	4,133.6	442,291.8
1855	.574	1,476.3	51,671.6	.662	5,083.5	427,013.3	.652	4,022.6	478,684.9
1856	.618	2,385.4	76,331.4	.709	3,604.7	262,321.7	.690	3,287.9	338,653.1
1857	.477	1,228.5	51,598.2	.703	3,886.3	264,267.0	.662	2,871.5	315,865.2
1858	.613	1,746.4	59,376.2	.709	4,130.0	161,069.1	.684	3,019.8	220,445.3
1859	.606	3,364.7	100,942.3	.718	3,251.0	152,776.1	.674	3,295.0	253,718.4
1860	.478	2,547.7	76,431.7	.725	4,342.1	312,634.3	.668	3,814.4	389,066.0

Year of Build	Total Iron and Wood Steamships								
	Paddle			Screw			Total		
	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity	Ratio of Net to Gross Tonnage	Transport Capacity per Ship	Total Transport Capacity
1823	.523	990.5	3,990.4	—	—	—	.523	990.5	3,990.4
1824	.500	793.2	3,172.8	—	—	—	.500	793.2	3,172.8
1825	.469	802.6	4,013.2	—	—	—	.469	802.6	4,013.2
1826	.526	795.8	8,754.2	—	—	—	.526	795.8	8,754.2
1827	.681	1,331.0	3,993.0	—	—	—	.681	1,331.0	3,993.0
1828	.447	1,780.0	534.0	—	—	—	.447	1,780.0	534.0
1829	.491	798.5	3,193.9	—	—	—	.491	798.5	3,193.9
1830	.589	781.5	3,907.6	—	—	—	.589	781.5	3,907.6
1831	.616	1,041.9	4,167.7	—	—	—	.616	1,041.9	4,167.7
1832	.610	1,031.2	8,249.7	—	—	—	.610	1,031.2	8,249.7
1833	.412	583.9	4,670.9	—	—	—	.412	583.9	4,670.9
1834	.581	1,374.1	23,359.8	—	—	—	.581	1,374.1	23,359.8
1835	.543	1,008.7	22,192.0	.778	3,275.2	3,275.2	.567	1,212.7	25,467.2
1836	.565	1,257.9	25,158.6	—	—	—	.564	1,257.8	25,158.6
1837	.578	1,208.5	39,879.2	—	—	—	.578	1,208.5	39,879.2
1838	.511	507.7	18,277.6	—	—	—	.511	527.2	18,977.6
1839	.537	819.6	17,211.1	—	—	—	.537	819.6	17,211.1
1840	.539	1,119.7	33,589.9	.524	8,809.5	17,619.0	.533	1,600.3	51,208.9
1841	.632	2,977.9	101,247.8	—	—	—	.632	2,977.9	101,247.8
1842	.567	1,711.5	54,769.5	.678	626.4	626.4	.569	1,678.7	55,395.9
1843	.512	626.5	20,674.0	—	—	—	.512	626.5	20,674.0
1844	.562	1,250.6	65,029.5	.637	604.4	1,813.3	.564	1,215.3	66,842.8
1845	.556	943.8	38,695.9	.651	1,090.2	3,270.7	.564	953.8	41,966.6
1846	.604	1,478.1	63,560.0	.712	2,248.9	24,737.7	.633	1,635.1	88,297.7
1847	.586	1,620.8	126,419.1	.723	2,253.5	13,520.8	.601	1,666.0	139,939.9
1848	.535	950.8	61,801.1	.556	2,382.1	14,292.5	.540	1,071.7	76,093.6
1849	.611	1,830.4	76,876.6	.692	3,610.4	28,882.8	.633	2,115.2	105,759.4
1850	.593	3,109.6	93,289.4	.693	2,100.0	29,399.9	.619	2,788.4	122,689.3
1851	.622	3,063.9	116,427.1	.671	4,556.7	72,907.6	.642	3,506.2	189,334.7
1852	.575	1,347.2	67,361.3	.669	6,154.5	135,399.7	.638	2,816.1	202,761.0
1853	.593	1,798.9	91,742.4	.678	4,825.6	221,977.9	.652	3,234.2	313,720.3
1854	.585	1,701.4	86,772.3	.697	4,655.9	395,747.9	.677	3,547.9	482,520.2
1855	.587	1,350.3	85,072.0	.663	4,977.1	428,026.7	.650	3,443.6	513,098.7
1856	.570	1,467.1	110,029.2	.709	3,615.8	263,953.4	.669	2,526.9	373,982.6
1857	.408	648.2	62,873.6	.702	3,639.6	265,691.1	.630	1,932.7	328,564.7
1858	.563	1,307.8	87,619.5	.707	3,758.0	161,592.2	.656	3,886.9	427,563.8
1859	.556	1,578.2	108,892.5	.718	3,023.6	160,249.6	.650	2,217.1	270,483.6
1860	.464	1,337.4	86,930.8	.724	3,928.9	314,308.8	.656	2,799.1	403,063.9

By 1865 steamships accounted for some 20 per cent of the total register tonnage, but Sir John Clapham thought it possible that the steamers did more transport work than all the sailing ships combined.[18] The argument applied (although to a lesser degree) to the late 1850's as well; the transport superiority of Britain's new steam fleet no doubt made important contributions to Britain's balance of payments.[19] To the extent that this was so, the rising trade deficit incurred from the 1850's onwards was offset more easily year by year because of freight earnings, especially since the proximity of coal and iron to the sea, together with British shipbuilding skills, brought a decisive shift in comparative advantage in shipping to Britain.[20]

Data exist which tend to support this view. It can be seen in A. H. Imlah's estimates of the current account shown in his pioneering work[21] on the British balance of payments. They show that the current account was especially strong (almost continuously in surplus from 1856 to 1874) in the third quarter of the 19th century, and that one of the main elements of this strength came from the sharp increase in net shipping earnings. If Clapham was correct, then the rapidly growing steam fleet (and its capacity to transport cargo, below) was no doubt an important contributor to Britain's commercial strength in the period.

By the late 1850's the development of the iron-screw steamer had made it already clear that Britain would gain greatly in her share of ocean freight earnings in the last half of the 19th century. The growth of the tonnage of iron-screw steamers in the British steam fleet shown in Table 363 underlines this point graphically. In terms of transport capacity, Britain's new steam fleet was even more impressive by 1860.

4. THE MEASUREMENT OF TRANSPORT CAPACITY

The customary measure of the cargo transporting capacity of a ship is its gross tonnage; the customary measure of the amount of shipping capacity built (or existing) in any year is the sum of the gross tonnages of the ships in question.

The adequacy of this measure of the amount of shipping is vitiated in the present case by the fact that important and rapid changes in design and construction of ships and engines took place during the period covered. Comparisons of the data for individual ships early and late in the period indicate that changes were made in hull design, evidenced by changes in the magnitudes and ratios of the linear dimensions of ships and in the changing composition of the fleet as between wood and iron hulls. Changes were made in engine design,[22]

[18] *Op. cit.*, p. 72.

[19] Although the data in table 367 do not include a breakdown between ships employed in river or coastal trade, and those employed in ocean trade, it should be borne in mind that many of the smaller steamers were used only in river or coastal trade. Again, it is also possible that many of these steamers were used mainly for passenger service. It would be unwise therefore, to *overemphasize* the contribution made by the steamers to the balance of payments.

[20] Largely at the expense of the American merchant fleet. Note the failure of American construction of steam and motor ships to continue to grow, in terms of tonnage, from 1853 to the 1880's. *Historical Statistics of the United States*, Washington, 1949, p. 211. Also the sharp decline after the 1850's (beginning then) in the growth of American tonnage employed in foreign trade leading to absolute decline beginning in 1862. *Ibid.*, pp. 207–8.

[21] "British Balance of Payments and Export of Capital 1816–1913," *Economic History Review* (Second Series), Vol. V, No. 2, 1952.

[22] See Graham's account, *Passim;* of changes in steam engines employed in ocean shipping. For an excellent account of the technical improvement of shipping during this period see Edgar C. Smith, *A Short History of Naval and Marine Engineering*, Cambridge, 1937, esp. pp. 94–186.

evidenced by systematic differences in the portion of gross tonnage occupied by engines and fuel for ships of given horsepower; changes were made in propulsion equipment, evidenced by the changing composition of the fleet constructed year by year in favor of screw propulsion at the expense of the paddle. In these circumstances a gross ton does not denote the same cargo transporting capacity for every ship. Thus, the changing composition from year to year of the fleet constructed means that an aggregate gross ton does not signify the same cargo transporting capacity in different years. In view of the rapid rate of technological advance in this period we are led to seek a more satisfactory measure of the amount of shipping constructed, where we view shipping as transportation capacity.

The transportation capacity of a ship may be considered its ability to move cargo a given distance in a given time. More specifically, we can consider the cargo transporting capacity of a ship to be:

$$\text{Number of cargo ton} - \text{miles per hour} = \begin{pmatrix}\text{net tonnage}\\ \text{of the ship}\end{pmatrix} \times \begin{pmatrix}\text{cruising speed}\\ \text{of the ship}\end{pmatrix}.$$

We shall call this quantity the *cargo transporting capacity*, or more briefly, the *transport capacity* of the ship.[23] By summing the transport capacities of the ships built in any year we obtain a measure of the transport capacity of the fleet built in that year, which we may call the aggregate transport capacity. The validity of this measure of shipping capacity is not affected by changes in the composition of the fleet from year to year or by changes in the design of ships, because the significant consequences for cargo carrying capacities of such changes are measured directly by this quantity. Thus, changes in hull design of ships produce changes in the space available for cargo, and in their speed; changes in engine design similarly produce changes in the space available for cargo and in speed; changes in design of propulsion equipment similarly affect cargo space and speed. (E.g., more efficient propulsion equipment makes it possible to reduce the size of engines and hence the space taken by engines and fuel without sacrifice of speed; or to increase speed without reducing cargo space.)

Computation of the transport capacity in the present case was impeded by the fact that the official data do not include the speeds of ships. They do, however, include the indicated horsepower of the engines and the linear dimensions

[23] A fully satisfactory measure of transport capacity would include a measure of that portion of port time (time spent in docking, loading and unloading) attributable to the ship. Our presumption is that port time, like sailing time, decreased as the technology of the steamship developed and spread into use. To the extent that this was so the transport capacities shown in Table 367 tend to understate the levels and the rate of increase over time of the actual capacity of the fleet to transport cargo. The presumption that port time would be shorter for steamers than for sail is supported by the following considerations.

Steamers represented substantial commitments of capital. Taylor estimates that "a steamship ordinarily cost three or four times as much as a sailing packet of comparable size." Taylor, George R., *The Transportation Revolution 1815–1860*. Rinehart and Co., Inc., New York, 1951, p. 119.

The large commitment of capital, the relatively smaller cargo space and the substantially larger proportion of port time to total time of transport for steamers would in the presence of the spur of competition, provide substantial inducements to reduce port time. Moreover the relatively rapid deterioration of iron steamers from rusting would provide an additional inducement in the same direction.

We have only isolated pieces of information on port time, and none of it separates that part due to port facilities from that part due to the ship. However, what data we have tends to support the view that port time was less for steamers than for sail. This is clear in the case of the coal trade, where the introduction of steam colliers encouraged the introduction of power equipment at docks to facilitate loading and unloading (see p. 378, 379 below).

of each ship. Because the speed of a ship of moderate speed depends to a first approximation on the power available to drive it and on the resistance due to skin friction with the water, it was possible on the basis of this information to estimate the speed of each ship by means of certain conventional formulas to be found in standard reference works on naval architecture and marine engineering. The speed of each ship was estimated by means of the following formulas:[24]

$$V = C \cdot \sqrt[3]{25{,}000 \frac{P}{S}} \tag{1}$$

$$S = L\{1.5H + .73B\} \tag{2}$$

In these formulas, P denotes horsepower, S, the wetted surface of the ship in square feet, V the speed, in nautical miles per hour, L the length of the ship in feet, H the mean molded draft (depth of hold) in feet, B the width in feet, and C a constant determined from data.

The transport capacity of the fleet could then be determined from the estimates made of the transport capacities of the individual ships using their estimated speeds.

Those considering using this measure of cargo-carrying capacity must settle for themselves the question whether it is preferable to use a measure correct in principle, but subject to an error of unknown magnitude analogous to an error of measurement, or to use an alternative measure, which is in principle unsatisfactory, but whose value is perhaps known more precisely.[25]

The Cunard Steamship Company was generous in making available the speeds of some twenty ships built during the period 1845–1861.

Trial computations for these ships led to the decision to divide the ships into two groups, those built prior to 1848, and those built after 1848. For those prior to 1848 the value $C = 1.11$ was used, for those after 1848, the value $C = 1.39$ was used. The following table shows the estimated and actual speeds and the error for the ships for which the speed is available.

The measures of transport capacity obtained with the aid of these formulas are given in Table 367. This table also shows the ratio of net to gross tonnage in each year (in effect a weighted average of the ratio of net to gross tonnage of each ship weighted by gross tonnage), and the transport capacity per ship in each year. These data enable us to compare the fleet measured by transport capacity with the fleet measured by gross tonnage.

---

[24] The formulas are taken from Taylor, D. W., *The Speed and Power of Ships; a Manual of Marine Propulsion*, 1933, pp. 171, 20–21 respectively.

Formula (1) is obtained from Kirk's Formula by taking the constant $k = 4$, the figure suggested by Kirk for ships with a high coefficient of fineness.

Formula (2) is obtained from Normand's formula by taking the block coefficient equal to .65, the value given by Hughes, Charles H., *Handbook of Ship Calculations Construction and Operations*, 3rd Edition, New York and London, 1942, p. 170. See also Machrow, Clement, and Woollard, Lloyd, *The Naval Architect's and Shipbuilder's Pocket-book of Formulae, Rules and Tables and Marine Engineer's and Surveyor's Handy Book of Reference*, Fourteenth Edition (Revised), pp. 175, 176.

Judgements involving the shapes of ships, based on examination of these data, on their linear dimensions and on such illustrations of early steamships as were available to the authors, were used as a basis for selecting the values of the two constants, k, and the block coefficient, used in these formulas. The arbitrary nature of this method of choice was partly offset by the introduction of the constant C, which was estimated from numerical data, see below.

[25] These hard alternatives may be softened by the fact that there is some information about the errors due to estimating speed. See Table 372 below.

TABLE 372

ACTUAL, ESTIMATED SPEED, AND ERROR FOR VARIOUS SHIPS BY YEAR OF BUILD

Ship	Years of Build	Actual Speed in miles per hour	Estimated Speed	Error
Cambria	1845	9.25	10.9	+1.65
America	1848	10.5	10.9	+ .40
Canada	1848	10.5	11.4	+ .90
Niagara	1848	10.5	11.3	+ .80
Asia	1850	12.25	11.5	− .75
Africa	1850	12.25	10.7	−1.55
British Queen	1851	10.0	10.8	+ .80
Delta	1853	9.0	9.1	+ .10
Melita	1853	9.5	9.1	− .40
Karnak	1853	9.5	10.8	+1.20
Emeu	1854	9.5	10.6	+1.00
Jura	1856	9.5	11.7	+2.10
Etna	1856	9.5	12.4	+2.80
Persia	1856	13.8	12.8	−1.00
Australasian	1857	12.5	9.5	−3.00
Olympus	1860	10.5	9.5	−1.00
Atlas	1860	10.5	9.6	− .90
Marathon	1860	10.5	9.6	− .90
Hecla	1860	10.5	9.6	− .90
Kedar	1861	10.5	9.5	−1.00

Source: Private Communications to the authors from Cunard Steamship Company, Ltd.

In Chart 373 we plot the logarithms of gross tonnage and transport capacity, respectively, against time. These series move closely together, except perhaps for the period 1848 to 1854.

Linear regressions of these series on time yield the least squares trend lines shown in Chart 373. It is seen from these that the original series have exponential trends, and that transport capacity increased during this period at a faster rate than did gross tonnage. The method of estimating the speeds of the ships affords grounds for the belief that the computed transport capacity may overstate the capacity of ships built early in the period and understate the capacity of those built late in the period; so that the rate of increase of transport capacity is lower than it should be.

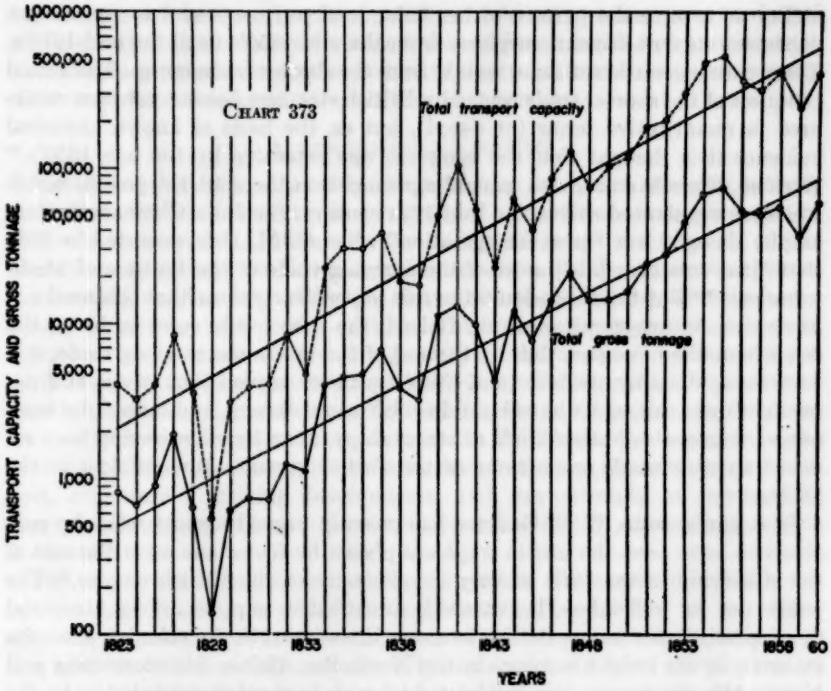
The equations of the trend lines are:[26]

$$\text{estimated gross tonnage} = 2.63 - .061 \times (\text{time})$$

$$\text{estimated transport capacity} = 3.26 - .067 \times (\text{time})$$

The data in Table 367 sharply amplify the preponderence of iron steamers in the total fleet; and also indicate some advantages of screw propulsion over paddle. By the 1850's iron steamers had outstripped wood both in total and in average transport capacity. Wood-screw ships did not consistently exceed

[26] The squares of the standard errors of estimate for these regression lines are .06 and .07 respectively.



wood-paddle steamers in transport capacity, but did in terms of the ratio of net to gross tonnage—indicating the larger amount of free cargo space which screw propulsion made available compared to paddle. It is notable that, in this respect, wood-screw ships were superior to iron paddle ships and even compared favorably to iron-screw steamers. Nevertheless, as is seen in Table 367, very few wood-screw steamers were placed under permanent registry in this period. This evidence supports the view that wood had disadvantages for ocean transport for which the compensation of screw propulsion and the advantages of long familiarity in construction were insufficient.

By the 1850's the highest total *and* average transport capacities had clearly passed over to iron-screw steamers. The sharp decline in both the total and average transport capacities of the wood ships after the mid-1850's probably indicates that the wood steamer was increasingly used only for river transport, and new wood steamers were no longer being purchased for employment in the ocean shipping routes even though a considerable number of wood-paddle steamers were still being placed under permanent registry (Table 363).

The transport capacity measures in Table 367 together with the data in Table 363 suggest some significant points about the role of the steam fleet compared to the sailing ships that still comprised the bulk of British merchant ships under British registry in the mid-19th century. As was argued above, the freight earnings from the new steam fleet doubtlessly added considerably to Britain's

ability to reverse the pattern of her balance of payments and to run almost persistent current-account surpluses from the mid-1850's until the mid-1870's. These earnings no doubt came mainly from the shorter trade routes. The actual conquest of the shorter trade routes by British steamers has not yet been measured in quantitative terms (or dated), but on the basis of known historical evidence it is thought that the conquest was extensive by the late 1850's.[27] Even so, there has not been general agreement on the actual degree to which steamers supplanted sail in the mid-19th century. Sir John Clapham, for example, thought it a "great exaggeration" when C. M. Palmer stated in 1864 that "iron steamers [did] most of the carrying trade of the Baltic and Mediterranean."[28] But the data adduced in this paper clearly show that this need not have been an exaggeration at all. Indeed, the shoe would seem to be on the other foot. We have seen that by the end of the 1850's the carrying trade, if it were being done by steamer (and Clapham never argued that it was still dependent upon sail) must have been done by iron steamers, and indeed, by iron-screw steamers—all other kinds of steam ships taken together having been reduced to very small proportions of the British steam merchant fleet in the 1850's.

In a similar vein, G. S. Graham has recently argued against what he considers to have been the undue emphasis placed by historians upon the role of the steamship in the 19th century development of the British empire.[29] The point may be well taken, but it needs quantitative support. If by "Imperial development" one means British overseas investment in the colonies, then the steamer, by its freight earnings in the North Sea, Baltic, Mediterranean and North Atlantic routes, was without doubt an important contributor to the strong balance of payments position which sustained British overseas investment in the last half of the century. Graham makes much of the inability of the steamers to compete with sail on the long routes to India, China, and Australia. But considering that more than half of British overseas trade was conducted within the possible steamer routes (Europe accounted for between 30 and 45 per cent, North America for another 10–30 per cent[30] and this cargo trade data does not include the very profitable emigrant passenger trade in which steamships participated), and considering that the carrying trade of Europe was within the sphere of British steamers, it would seem that it must be shown in quantitative terms that freight earnings from the long-haul cargos of sail were greater than steamer earnings if Graham's point is to be substantiated.[31] Until such evidence is forthcoming it would not be wise to write off the "traditional" emphasis on the role of steam. The data in Tables 363 and 367 above show that by the end of the 1850's Britain's steam merchant fleet

[27] E.g. Graham, *op. cit.*, p. 74. Throughout this account there is scattered evidence on this problem. G. R. Taylor states that " . . . steamships had, during the fifties, been able to take over the best-paying part of the ocean trade." He believes that in total world trade, by 1860, the sailing ship still prevailed "very largely." *The Transportation Revolution*, New York, 1951, p. 119.

[28] *Op. cit.*, p. 71.

[29] *Op. cit.*, pp. 74–75.

[30] See Werner Schlote, *British Overseas Trade from 1700 to the 1930's*, Oxford, 1952, Tables 18, 19, pp. 156–159.

[31] And this may be conceivable if data for freight earnings can be uncovered, since bulk cargos of the kind carried by sail may well have earned more freights (in aggregate) than manufactured goods and passengers. The authors are indebted to Douglass C. North for this observation based upon his as yet unpublished investigation of 19th century ocean freight rates.

was not the wooden assortment of paddle wheelers that many have envisaged (above), but that it was mostly iron-screw, of recent build and growing with increasingly greater transport capacity. It was a commercial weapon worthy of the importance assigned to it by its contemporaries.

5. THE HISTORICAL SETTING

New techniques have only rarely affected major changes in economic activity with what might be called startling rapidity. The case of the iron-screw steamship was no exception. Indeed, by the mid-1850's when the triumph of this kind of ocean transport in the steamship routes was obvious to contemporaries (and visible on the ways of British shipyards) the steamship already had a history stretching back half a century. The iron-screw steamer had been in use for at least two decades, and a great and colorful age of commercial sail on the longest ocean routes still lay in the future.[32] Yet, given the history of experimentation and growth in earlier decades, the rapid increase in the number and tonnage of iron-screw steamers in the British merchant marine in the 1850's together with the resulting displacement of other forms of ocean transport, constitute a striking development; and was certainly an outstanding example of the force of innovation in the record of British economic growth in the period.

The accelerated introduction of the iron-screw steamer in the 1850's can best be understood in its own historical context. The demand for iron-screw steamers depended mainly upon the needs of overseas trade. A major innovation such as this one might successfully make inroads upon the markets of older and less efficient kinds of shipping because of cost advantages, but a rapid expansion of the *total* market would considerably ease the problem of entry, especially if existing shipping capacity were very nearly utilized. This would presuppose a strong growth of foreign trade, which in turn would, *ceteris paribus*, be related to a sharp rise in domestic and foreign economic activity. Such conditions prevailed in the decade of the 1850's. Moreover, the rise of a competing form of transportation, the railway, threatened the coastal trade in coal and virtually forced the quick adoption of the iron-screw steamer in the coastal transport of coal. Here was a process familiar in economic history—innovation giving rise to innovation on a grand scale.[33]

The British economy evolved slowly from its mainly agricultural base toward an "industry state" in the half-century following the end of the French wars in 1815. The growth of industry and trade was uneven, but with the steady increase in population, internal improvements in industrial techniques, transportation, and communications together with the increasing movement of laborers from the countryside into the industrial towns,[34] the transformation

---

[32] Gerald S. Graham, *op. cit.*, p. 75.

[33] For an extensive discussion see, for example, Paul Mantoux's discussion of the effects of Kay's flying shuttle on the textile industry, *The Industrial Revolution in the Eighteenth Century*, London, 1949, pp. 211–3; and Karl Marx, *Capital*, Vol. I (Glaisher, ed.) London, 1918, pp. 379–82, on the effects upon invention of disparate increases in productivity in related industries caused by innovations; e.g., . . . it was only during the decade preceding 1866, that the construction of railways and ocean steamers on a stupendous scale called into existence the cyclopean machines now employed in the construction of prime movers." p. 380.

[34] This movement was at its height in the 1840's; see A. K. Cairncross, "Internal Migration in Victorian England," *Manchester School*, 1949, vol. xvii, p. 70.

became ever more perceptible. The outflow of British people (and other Europeans) to America and to the colonies[35] insured a growth of British overseas trade in industrial products. This growth was augmented by commercial development elsewhere and by the ever-growing interdependence of commercial and financial activity on a world scale. In a very real sense, trade engendered more trade. At the center of this complex of markets the growing British industrial economy became increasingly dependent upon overseas supplies of raw materials, and (by the mid-19th century) of food.

The effects upon Britain's overseas trade are to be seen in Chart 365. Even before the great upsurge of economic activity that characterized the decade of the 1850's, the movement of British foreign trade was strongly upward in volume terms. This no doubt provided enough of a growth element in the ocean-freight market to buoy up the expectations of investors in steamers. Thus in the late 1830's and twice in the 1840's there were fairly marked increases in the construction of steamships, either wood paddle or iron paddle, to be added to the British merchant fleet. But as can be seen in Table 363 it was in the years immediately after the Californian and Australian gold discoveries and during the Crimean War that the character of steamship construction changed radically and the iron-screw steamer clearly displaced wood and iron paddle steamships in United Kingdom shipyards.[36]

While the rapid growth of iron-screw steamer construction is evident as early as 1849, the years 1850 to 1855 saw a spectacular increase in their construction and the introduction of the iron-screw steamer was decisive. The boom in overseas trade (especially raw-material imports and also exports of British commodities to America and Australia up to 1854, together with a sharp rise in exports to the Continent and to the Middle East in 1855) followed by military transport requirements while trade remained at very high levels, led to a strong demand for new shipping capacity, and during the war, especially for steamers.[37]

The rise in ocean freights even before the outbreak of War with Russia in 1854 was very marked and spurred the efforts of shipbuilders and ship speculators. From the first quarter of 1852 to the third quarter of 1853 average cargo freight rates from British ports to New York rose from 15s to 28s per ton. During the same period, freights on (measurement) goods to Calcutta rose

---

[35] From 1845 to 1850 alone roughly 1.3 million people emigrated from the United Kingdom. *Statistical Abstract.*

[36] The change to iron was marked in sailing vessels as well. "As regards iron-built sailing vessels, the most remarkable feature of our trade is the very increasing favor they are now growing into, and which are now occupying the builders both here [Liverpool], on the Clyde, Newcastle and elsewhere to an unprecedented extent." *Circular to Bankers*, January 8, 1853. By the middle 1850's a shrewd and observant contemporary, Thomas Tooke, without distinguishing between kinds of iron ships, saw clearly the change that had taken place. "The number of Iron Vessels now owned by this country is very great, and it is rapidly on the increase. On the Clyde it is a rare sight to see a ship built of wood. The Tyne-built are, also, almost all of iron. At Liverpool also, iron is becoming much used; and even at Sunderland, where they have been so long accustomed to wooden vessels, some of the leading builders are directing their attention to iron." *History of Prices*, London, 1857, Vol. V, p. 331.

[37] The striking rapidity and magnitude of the response of the shipbuilders to the stimulus of increased demand was partly made possible by the flexibility of the British financial system of the period. Ships were financed on the basis of long-dated bills of exchange which were accepted and discounted like ordinary trade bills in the London discount market. The system was largely discredited in 1866 when the great discount house of Overend and Gurney failed and their underwriting, by this means, of the notorious ship speculator, Stephanos Xenos, became known. In *Accounts and Papers*, 1861, LVIII (371) Xenos appears prominently from the mid-1850's onwards in the London registrations.

from 22s to 42s per ton, and to Melbourne, from 42s to 85s per ton.[38] Such increases of as much as 100 per cent reflect the shortage of shipping capacity that existed in the wake of the trade expansion of the early 1850's. But when war came early in 1854 freights rose even higher in response to government demand for troop transports to the Black Sea.

Immediately following the outbreak of war government transport demand had been mainly for sail, but several large steamers were chartered from the principal steamship companies and these ships soon proved their superiority. They were mostly of large size[39] and made the voyage from Portsmouth to Constantinople in from 12 to 15 days. The same trip for sail was an estimated 40 to 50 days.[40] Accordingly government charter demand swung over to steamers: there were 31,900 tons of steam under government charter by June 1855 compared to only 15,800 tons of sail; a year earlier these figures had been virtually the reverse, 14,000 tons of steam and 39,800 tons of sail.[41] By the end of 1854 the government was paying 55s to 60s per register ton for steamers; 25s to 30s had been the charter price at the beginning of the war.[42]

There developed a shortage of steamers in almost every route when government demand was met.

> " . . . though every line of the trade has been drained, even to the derangement of our postal services, sufficient steamers cannot be found.[43]

By late 1854 it appeared that British shipping capacity was strained virtually to the limit. Wages of shipbuilders and seamen by 1854 had risen 50 per cent or more compared to 1850.[44] The rise in shipbuilders' wages reflected both private and government demand for newly constructed ships. Of the 354 steamships added to the British Navy from 1848 to 1859, 284 were purchased in the three years 1854–56.[45] As can be seen in Table 363 the increased tonnage of steamer construction (and especially of iron-screw steamers) in the early 1850's reached boom proportions during the Crimean War years. During that period the market seemed to warrant a great expansion of steamship construction—and especially of iron-screw steamers.[46] But the Crimean War shipbuilding boom largely accounted for the depression in shipbuilding after 1855; excess capacity had been created, and especially was this the case when the war ended.

---

[38] Tooke, *op. cit.*, p. 319.

[39] Averaging well over 1,000 gross tons, the largest, the *Himalaya*, chartered from the Penninsular and Oriental Steam Navigation Company, had a gross tonnage of 3,550. *Ibid.*, p. 320.

[40] *Ibid.*

[41] *Ibid.*

[42] *Ibid.*, p. 321.

[43] *Times*, January 4, 1855.

[44] The peak of wage rates reached during the Crimean War represented the highest levels of the whole period 1848–1860; wages declined from 1855–56 onwards. See Bowley, A. L. and Wood, G. S., "Statistics of Wages in the United Kingdom . . . Engineering and Shipbuilding," *Journal of the Royal Statistical Society*, 1906, Vol. LXIX, p. 518; for wages of seamen, see *Accounts and Papers*, 1867, LXIV (158).

[45] Accounts and Papers, 1859 (Sess. 1), XV (180). For total vessels ordered by the Navy and launched in this period see *Accounts and Papers*, 1860, VIII (545), *Report of the Select Committee on Navy Guns and Mortar Boats*, Appendix 3.

[46] Moreover, the steamers constructed were mostly of large size suitable for overseas trade, while the coasta trade was neglected. Steam tonnage employed in coastal trade fell from 85,500 in 1853 to 54,000 in 1854, and was only 57,400 in 1855. *Times*, January 4, 1955; *Statistical Abstract*. This no doubt helped to shift the transport of coal to the railways during the War years and especially in 1855. See below pp. 378–380, especially Table 379.

As can be seen in Table 363, the decline in shipbuilding in the years immediately following 1855 was sharp—and this in spite of the unprecedented levels of trade ruling in 1856–1860. Here the yards specializing in the construction of iron-screw steamers were evidently not protected by the technological superiority of their product; they were hit as badly by the excess shipping capacity built during the Crimean War as were those shipyards still specializing in sailing ships.[47] However, even though new construction of iron-screw steamers was drastically reduced, the superiority of these ships over sail was becoming increasingly evident and the war-built steamers were evidently capturing the shorter trade routes as early as 1856.

> "Surplus steamers are now finding their way, to the advantage of all concerned, into trades formerly carried on by sailing vessels."[48]

The Baltic routes especially felt the impact of the Crimean War iron-screw tonnage. On these routes the fuel problem was less of a handicap than on other routes and by 1856 the intrusion of steamers was a notable feature of the Baltic trade. In 1853, 26 steamers had reportedly made 79 voyages from British to Baltic ports; in 1856 the number of steamers in that trade had risen to 136 and the number of voyages to 334.[49] Even as early as the beginning of 1857 some shipowners realized that steam was about to overtake sail in the shorter routes.[50] By the end of 1858 it was noted that steamers had succeeded in the Mediterranean trade.[51] In the coastal trade railway competition had already assured the introduction of steam.

The construction of trunk lines associated with the railway "mania" of the 1840's together with the proliferation of branch lines during the 1850's provided Britain with a vastly improved land transportation base for industrial expansion. But the new railways also broke the bottleneck in coal transport, a field that had formerly been the preserve of the coastal-shipping companies, and provided strong competition for the freight earnings of coal haulage. Development of a rapid internal transportation system capable of handling coal meant that the old wind-driven colliers faced a kind of competition they had never known before. Moreover, coal-producing areas not served by direct rail connections with London were faced with potential losses of a good part of their market in the metropolis and were forced to find a means of retaliation. Iron-screw colliers provided the means of competing.

From 1850 to 1860 coal transported into metropolitan London rose from 3,639,000 tons to 5,070,000 tons, an increase of 1,431,000 tons. During the same period, coal brought into the London district by inland navigation and railways rose from a mere 85,000 tons to 1,497,000 tons, or by 1,412,000 tons, almost the whole of the increase in the total.[52] By 1860 coal brought into London by coastal shipping stood at 3,573,000 tons, a mere 20,000 tons above the 1850

[47] The shipowners were well aware of their predicament. "In estimating the prospects for the current year, we cannot hide from ourselves that we have far too much tonnage afloat for the purposes of our legitimate trade." *Times*, January 3, 1857.

[48] Tooke, *op. cit.*, Vol. V, p. 331.

[49] *Times*, January 3, 1857.

[50] "In our opinion the time is rapidly approaching when for short voyages steam will be as a general rule substituted for wind." *Ibid.*

[51] *Ibid.*, January 3, 1859.

[52] *Report of the Commissioners Appointed to Inquire into the Several Matters Relating to Coal in the United Kingdom, Accounts and Papers*, 1871, XVIII (C. 435-II), Appendix to Report of Committee E., p. 975.

figure, but in 1855 coal brought into London by coastal shipping had fallen to just above 3,000,000 tons; at the same time inland transport had risen by over one million tons to 1,161,000 tons.[53] From 1855 to 1860, as is apparent from these figures, the coastal trade regained some of its earlier losses to the railways; coal brought in by coastal shipping increased more than that brought in by inland transportation. This change in the fortunes of the coastal trade was due almost entirely to the introduction of iron-screw colliers.

The Northern shippers and coal-owners had been fairly quick to perceive the advantages of iron-screw colliers in spite of a suggestion to the contrary implicit in the volume data shown above. As early as 1844 interest was aroused in London on the arrival of a Newcastle collier ". . . a fine looking vessel built of iron and fitted with a screw propeller, working an engine of 20 horse power."[54] By the 1850's the beginning of the shift of the English iron industry's center to the North Riding of Yorkshire, following the successful processing of the Cleveland "yellow band" iron ores, produced a powerful group of new industrialists who quickly adopted the iron-screw steamer. One of these men, C. M. Palmer, a Northern coal and iron owner as well as a shipbuilder, launched the "John Bowes" in 1852. This ship was an iron-screw collier built to carry 650 tons of coal at 9 miles per hour. On its initial voyage, the "John Bowes" went from Newcastle to the Thames in 48 hours, unloaded in 24 hours, and was back in Newcastle in another 48 hours, having done a job in five days that " . . . would have taken two average sized sailing colliers upwards of a month to accomplish."[55] As can be seen in Table 379, these results were not lost on Northern shipowners.

TABLE 379

NUMBER OF CARGOES AND TONS OF COAL CARRIED BY IRON-SCREW COLLIER, NEWCASTLE TO LONDON BY YEARS FROM 1852 TO 1862

Year	Number of Cargoes	Tons of Coal
1852	17	9,483
1853	123	69,934
1854	345	199,974
1855	174	85,584
1856	413	238,597
1857	977	547,099
1858	1,127	599,527
1859	899	544,614
1860	1,069	672,476
1861	1,299	851,991
1862	1,427	929,825

Source: J. S. Jeans, *op. cit.*, p. 75.

The hundred-fold increase in coal shipments by iron-screw collier in a matter of a decade was not only a consequence of increased production and use of these steamers, but was also a result of rapid technical improvements in such ships.

[53] *Ibid.*
[54] Harry Scrivenor, *History of the Iron Trade*, London, 1854, p. 311.
[55] J. S. Jeans, *Notes on Northern Industries*, London, 1879, pp. 74–5.

In 1857 an announcement appeared that the "William Cory" was ready for service between Cardiff and London. The explicitly stated purpose for building this ship was to compete with railway transportation. The "William Cory" was designed to carry 1,500 tons of coal—more than twice the capacity of the "John Bowes." The "William Cory" was also fitted out to take on water as ballast for its homeward voyage and the Victoria Docks had been supplied with hydraulic equipment that would take off the entire cargo in 16 hours.[56] Given the pace of technological advance implied by these two illustrations,[57] the data in Table 379 are easily enough understood. Although the iron-screw collier might have developed rapidly enough on its own, the vigorous railway competition of the 1850's accelerated the introduction of this new form of ocean transport.

The growth of the new iron-steam fleet in this period had broad economic implications. The construction of the new iron-steam fleet (and also the adoption of iron for sailing ships) provided an additional source of demand for iron and underpinned expanded investment and output in the iron industry and in the closely connected coal industry in the 1850's.[58] In this respect the demand effects of iron-ship construction on the iron industry were similar to those observed by Clapham in the period 1873–1883.[59] The shift of iron production toward the North East during the 1850's was thus no doubt given an additional stimulus by shipbuilding booms in the shipyards of the Tyne and Tees, and in Glasgow and other iron ports in the area. The number of steamer registrations in the North-East ports in fact rose rapidly from 1851 onwards.[60]

The trends in transport capacity and gross tonnage added to the British steam fleet 1823–1860 (above p. 372 and Chart 373) are of interest in connection with certain theories of economic development. From the time of Malthus theorists have speculated on the possibilities of exponential growth, and recently attention has again turned to growth models characterized by exponential growth.[61] If these models can be regarded as referring to average behavior of output, then the observed growth of transport capacity (Chart 373) is consistent with the hypothesis of exponential growth for the period 1823–1860.

Converting the regression equation for transport capacity into the form of an exponential to the base "e" we have:

$$\text{Estimated transport capacity in year } t = e^{(.154t+7.5)}$$

Transport capacity thus grew on the average at the rate of 15 per cent per year during the period observed. It is beyond the scope of the present paper to determine to what extent the growth of transport capacity is representative of over-all economic growth in the period. But the transport capacity data cer-

[56] *Times*, July 16, 1857.

[57] A further discussion about the rate of obsolescence in these early steamships is given above.

[58] A brief discussion of iron and coal output and investment in the 1850's will be found in J. R. T. Hughes, "The Commercial Crisis of 1857," *Oxford Economic Papers* (New Series), Vol. 8, No. 2, June, 1956, pp. 196–7, 204–5. For Scotland alone, R. H. Cambell, "Fluctuations in Stocks," *Oxford Economic Papers* (New Series), Vol. 9, No. 1, February, 1957, p. 43.

[59] *Op. cit.*, p. 61.

[60] With the exception of Sunderland.

[61] See for example; von Neuman, J., "AModel of General Economic Equilibrium," *Review of Economic Studies*, Vol. 13, No. 1, 1945–46, pp. 1–9; Solow, R. M., and Samuelson, P. A., "Balanced Growth under Constant Returns to Scale." *Econometrica*, Vol. 21, No. 3, July 1953, pp. 412–24.

tainly help to raise the possibility of exponential growth out of the realm of pure hypothesis; and a cursory examination made by the present authors of aggregate data for the British economy in the first half of the 19th century suggests that further research in this area will yield high returns to interested scholars.

The historical background of the boom in investment in steamers in the late 1840's and in the 1850's can add considerably to our understanding of the rapid expansion of this means of ocean transport in the period and contributes to our understanding of British economic growth in the period. But there remain many unanswered questions; certainly until adequate quantitative evidence on the actual competitive performance of these ships in the ocean-freight markets is forthcoming, the development of the ocean-going steamer in this early period will not be fully understood. What were the actual cost incentives of iron-screw steamers compared to other types in this period? How much longer than wood ships did the iron steamers remain in service? How long did the growth patterns indicated in Tables 363 and 367 above continue after 1860? When did the new steel ships begin to displace iron in the new steam fleet? These are questions which ought to be profitable areas for further investigation.

6. SUMMARY

Beginning in the late 1840's the growth of the British steam fleet began to accelerate markedly and this fleet was mainly built of iron. As early as 1844 the gross tonnage of iron steamers added to the fleet exceeded wood. By 1851 the gross tonnage of iron-screw steamers alone exceeded all wood steamers. Hence the "general" adoption of iron for steamships certainly began in the early 1840's, or from 10 to 20 years earlier than has customarily been believed. Some 60 per cent of the gross tonnage of all steamers added to the fleet in 1823–1860 was added in 1854–60, 85 per cent of this gross tonnage was iron, and about 70 per cent was iron-screw. Given the high rate of obsolescence of these early steamers, it has been argued that by 1860 the British steam merchant fleet was mostly iron-screw and less than a decade old. It was shown that due to technological improvements the rate of growth of the transport capacity of this fleet was even faster than the gross tonnage, and that iron-screw ships were predominant by these measures, and were the most efficient in terms of net cargo-carrying capacity. It has been argued that, in the light of these findings, historians have probably underrated the strength of the British steam fleet which existed by the early 1860's, and that the claims of contemporaries, thought by recent historians to have been exaggerated, were probably justified. At least there is little quantitative evidence to the contrary. Finally, it has been shown that the rapid development of the steam merchant fleet in the 1850's was associated with strong economic expansion both within Britain and in her overseas trade. The growth of the steam fleet was further buttressed by the Crimean-War transport demand and was given an additional stimulus by the growth of railway competition for the coastal trade in coal. Hence a favorable economic environment prevailed in Britain for the accelerated introduction of this radical innovation in ocean transport.

THE USE OF RANDOM WORK SAMPLING FOR COST ANALYSIS AND CONTROL*

A. C. Rosander, H. E. Guterman, and A. J. McKeon
Internal Revenue Service, Washington, D. C.

Random time sampling of activities (work sampling) is successfully applied to everyone in a division of 350 employees for the purpose of budget planning and cost control. To insure this success several technical and administrative problems had to be solved, especially those relating to the design and management of the sample. An appropriate model is described together with methods of sampling, estimating, and calculating sampling errors. Bias is reduced by using an employee instruction booklet, anonymous reporting, and by classifying and coding activities. Cost of the plan is reduced by using a stratified cluster sample, a check-type data sheet, a simplified method of estimation, and by calling random times by telephone.

PURPOSE

The purpose of this paper is to show how probability sampling can be applied to work and nonwork activities in order to obtain data for cost analysis and cost control. Methods are described which show how estimates of cost can be made directly from the sample. This procedure goes beyond the usual practice followed in ratio delay studies or work sampling where estimates are limited to percentages. Techniques applicable to office or service operations are described in detail and illustrated by an actual study. Several unique factors not heretofore used in work sampling are described in detail; one of these is estimating directly from the sample various costs and the sampling errors applicable to these costs. Methods of probability sampling and estimating are described which reduce the amount of work required, facilitate processing, and speed up the release of significant cost data to management.

HISTORY

The present technique of work sampling, also known as ratio delay study, seems to have had its origin in the early 1930's when L. H. C. Tippett [7], the English statistician, applied it to machines and machine operators in a textile mill. Tippett used what he called "A Snap Reading Method" to obtain at random intervals observations about machine operators for the purpose of estimating the per cent of time machines were idle, the percentage composition of reasons for delay or down time, per cent of time workers spent on each of their duties and the percentage distribution of production so as to enable overhead cost to be distributed among several products.

Tippett mentions four merits of work sampling: (1) it costs much less than the usual production study, (2) the method is more acceptable to workers and

* This paper was prepared and the project described herein was carried out under the general supervision of E. J. Engquist, Jr., Director of the Statistics Division.

labor unions because no stop watches are used, as in time and motion study, (3) better data may be obtained because, while workers may change their activities when they discover a production study is being made, they cannot do this so easily when work sampling is used, and (4) when work sampling is used, major sources of variation, as well as random sampling error, can be isolated.

Tippett made tests to determine whether or not the percentages derived by the sampling method were distributed as the binomial distribution. He found that this was true in some textile factory operations but warned that due to differences which might arise between time, machines, and operators the actual sampling error could exceed that given by the formula for the standard deviation of a proportion distributed as a binomial distribution where the proportion P is constant: $\sigma = \sqrt{PQ/n}$. He pointed out that relatively large samples were required—usually 500 or more.

Some years later Morrow applied Tippett's method to a number of industries and in general came to the same conclusions that Tippett did [5]. Morrow called this method "ratio delay study." He pointed out that the method can be used to estimate percentage of time consumed in delay, percentage of time spent on various machine and nonmachine operations, and percentage of time machines were working and not working. He pointed out also that the observations should be distributed over all hours of the day and week, that the method provides a way of observing and evaluating operations of the department as a whole, and that the observations may be interrupted at any time without affecting the study.

Waddell, editor of *Factory Management and Maintenance*, introduced the term "work sampling" in order to emphasize the fact that Tippett's technique need not be limited to the study of factors of delay but could be applied to the study of all work activities [3].

Brisley described how the method of random work sampling can be illustrated and validated by drawing random time samples from a moving picture of a person at work. The estimated percentages of time derived from these samples were compared with true percentages calculated from the film [3]. He applied the method of work sampling to nurses, professional engineers, and executives for the purpose of determining among other things how much time was spent on regular duties and how much on duties someone else should perform.

In 1952 Abruzzi reviewed the problem of delay factors and described how the quality control chart can be used in connection with Tippett's method in order to establish whether or not the pattern of delay is stable over time [1]. He reported that at one plant there were no significant differences in the delay percentages within days and between days where the period covered was 7 days. This finding is similar to that reported by Tippett relative to delay percentages found in certain cotton mill operations in England.

Barnes [2] compiled a wide variety of papers illustrating the applications of work sampling to the problems of industry. He described how the technique can be used not only to estimate relative frequency of idle status, sources of delay, and components of work operations but also to obtain data which can form the basis for wage incentives, work measurement, and performance rating.

A review of the foregoing papers and other papers in this field reveals a number of important points. In many instances detailed information is lacking as to how randomness is obtained in sampling machines and operators. No distinction is drawn between an individual sampling unit and a group or cluster sampling unit, with the result that inappropriate formulas are sometimes used to estimate sampling errors and sample sizes. Information is lacking relative to the most efficient sample design to use in any given situation. There is little or no information relative to stability of percentages over long periods of time, such as several months. This latter aspect is related to the problem of how best to distribute the sample over time. For example, in one case about 2300 observations were taken in 3 days, and 5000 observations in 5 days. If the variation between weeks is large compared with the variation within days, then the selection of large samples each day for a few days is inefficient because it not only eliminates week-to-week variation but it may give an inadequate measure of variation between days—types of variation which should appear in the estimates. It may be better to take fewer samples per day and include more days and weeks.

This review also shows that random work sampling may be preferred to other methods of cost analysis such as production studies, ticket systems, and test runs where each is used for the same purpose. In the production study continuous daily records are kept for each worker and machine showing each activity and the time spent on it. This record keeping is time consuming, costly, may be done carelessly, and may be biased if limited to short lengths of time. Furthermore, as Tippett pointed out, workers may deliberately change their work habits while such a study is in progress thereby introducing bias into the data. Morrow [5] reported that work sampling required only about one third of the time required to obtain similar data by the usual production studies. Barnes [2] cited a study where the time required for a work sample was less than 5 per cent of that required by the usual time study, yet the data were as good.

Most of the applications referred to above were made in connection with factory operations, and for the purpose of estimating percentage of work time lost due to major sources of delay or trouble. That the technique can be applied to other problems and to nonfactory work, is illustrated by the example described below.

THE ORGANIZATION AND THE PROBLEM

Random work sampling is being applied to the work of the Statistics Division in the Internal Revenue Service for the purpose of cost analysis and control. This Division consists of about 350 employees divided into two main branches—one devoted to planning and the other to operations. There are about 50 employees in planning and administration; the remainder are in operations. The Division is divided into a total of 24 different sections and offices.

The Division is a service agency providing professional service relative to the planning and approval of projects of a statistical nature, as well as providing large scale tabulation and computation services. This means that a wide variety

of jobs is required including those in the clerical, professional, and administrative fields. There are messengers, mail clerks, coding clerks, key punch operators, tabulation machine operators, typists, statistical clerks, and secretaries. There are electronic machine programmers, analytical statisticians, mathematical statisticians, and economists. There are supervisors, chiefs of sections and branches, and administrative and executive officers.

It was planned to collect several kinds of data:

(1) time and cost associated with major functions, such as program and analysis;
(2) time and cost associated with major programs, such as management information reports;
(3) time and cost associated with special services which are requested from outside the Division; and
(4) time spent by a section, an office, and the Division, on nonwork activities (a) due to production factors such as no work assigned, or (b) due to other factors such as leave, coffee breaks, and personal needs.

These data would make it possible for management to do the following:

(1) plan a sound and realistic budget for the entire Division;
(2) determine month-to-month fluctuations in costs;
(3) determine the additional resources that would be required by an expansion of function or program;
(4) determine the resources that would be released by restricting a function or program; and
(5) isolate weak and wasteful spots in the organization.

With these data available management would be in a better position to plan, to appraise, and to take remedial action. It was not intended that the data obtained would be used to take action relative to individuals since the data records were to be anonymous, but instead the data obtained for a section or office would be used by the supervisor only on a group basis.

THE SAMPLE DESIGN

The test run. Once the use of work sampling in principle was approved by the director, it was necessary to develop a sample design and a method of collection which would be feasible under the various limiting conditions which had to be met.

In view of the almost complete absence of knowledge relative to the characteristics which were to be estimated, and the kind of data collecting system which would work, a pilot or test run was made during one week of five consecutive days in June 1955.

During that week eight random time samples were taken daily of the work of every person at work in the Division. Random times were called by units or offices, so that the unit of sampling was a cluster of persons [4]. The time unit randomly selected was the minute, and a table of random numbers [6] was used to select the sample minutes from a total of 480 minutes for those units or offices where the lunch hour was excluded from the sample, and 510 minutes for those units or offices where the lunch hour was included in the sample.

The materials prepared and distributed for this test run included data record sheets, an instruction booklet, and a coded list of projects. Verbal instructions were given to all employees prior to the start of the test run.

Five higher grade clerks were used part time in order to inform units or offices of the random times and to collect and review data records. This latter step was necessary to insure proper understanding of instructions and to account for all employees.

Results of the test run. The results from the test run were so important that they more than justified the extra work involved. The major findings, which were incorporated into the final plan, were as follows:

1. 8 random times daily were unnecessary; no more than one per day was required. Table 386 illustrates the homogeneity of an activity since a chi-square test shows no difference between morning and afternoon. Hence in a case like this, it is unnecessary to sample both morning and afternoon.
2. Variability was much greater among personnel in the Program Branch than among those in the Operations Branch, so different rates of sampling were called for.
3. More detailed instructions to employees were called for. Consequently, an enlarged booklet of instructions was prepared including code numbers to be applied to activities not covered by the list of projects. In addition, employees were to be given the code number of every job to which they were assigned.
4. Every supervisor was asked to review the data sheets for his unit to insure that (1) one data sheet was obtained from every person for every time called, and (2) each sheet contained no gross errors such as wrong salary grade, wrong organizational unit, or wrong type of activity.

TABLE 386

FREQUENCY OF OCCURRENCE OF ACTIVITY A IN RANDOM TIME SAMPLES COVERING 5 SUCCESSIVE DAYS

Organizational Unit	A.M.	P.M.	Total
1	74	87	161
2	51	56	107
3	22	22	44
4	11	21	32
Total	158	186	344
Work time division	50%	50%	100%
Expected frequency	172	172	344

$$\text{Chi-square} = \frac{(158-172)^2}{172} + \frac{(186-172)^2}{172} = 2.28$$

THE FINAL SAMPLE DESIGN

As a result of the test run, the sample was re-designed so as to give management adequate information on a timely basis. This sample was designed to give monthly estimates of percentages and costs for broad types of activities and major projects, for the Division as a whole. It was not designed to give adequate data for specific projects nor for small administrative units on a monthly basis. However, over a period of several months adequate estimates could be obtained for major projects and activities for each of the 24 administrative units.

The total number of clusters drawn at random weekly is 3×5 units$+2\times19$ units, or 53. This means 212 clusters every 4 weeks or 2,756 clusters during a year. These sample sizes of 212 and 2,756 are sufficient in view of the stability of the characteristics being estimated and the sampling errors which can be tolerated for management purposes.

Two strata were introduced: the Program Branch with about 50 employees and the Operations Branch with about 300 employees. Three random observations are taken every five days for the first group, and two random observations are taken every five days for the second group. This plan gives about 3,000 observations every four weeks for the entire Division. ($3\times4\times50+2\times4\times300=3{,}000$.)

Stratification is used because of the greater variability of the activities in the Program Branch. In this latter group which includes the higher level professional and managerial personnel, the projects and activities engaged in vary greatly. On the other hand in the Operating Branch large numbers of employees may be engaged on one project for weeks at a time.

The random times are assigned at random to the days within each five day period, and only one random time is selected per day for any administrative unit. In other words the two or three random times cannot occur during one day—only one random time is assigned to any one day.

The rate of sampling for the first group is three divided by the total number of minutes in 5 days, or

$$\frac{3}{510\times5}=\frac{1}{850}.$$

The rate of sampling for the second group is two divided by the total number of minutes in 5 days or

$$\frac{2}{2550}=\frac{1}{1275}.$$

The total number of minutes in a day is $480+30$ or 510, since the lunch hour is included in the sample. Since the sample unit is a minute, we can think of the population as finite and express the sampling rates as given above.

As a result of the test run, the lunch hour of every employee is included in the time subject to random sampling. It is easier to do this since not all employees have fixed lunch hours while all of those who do have fixed lunch hours do not have them at the same time. Observations which fall in the lunch hour

can be excluded from the tabulations, or if desired, they can be used to estimate whether employees in the Division as a whole are taking more than the allotted time.

For each random time every person in the organizational unit fills out a data sheet. Hence the sampling unit subject to random selection is a *cluster of persons*, not an individual. This means that while the formula for the sampling error of a proportion derived from a binomial distribution $\sqrt{PQ/n}$ is applicable to the repeated sampling of individual sampling units, where a person is engaged in activity A or not, it is not applicable to group or cluster sampling units. The appropriate expression for the sampling error for the latter type of sampling is described and illustrated below.

Minutes rather than individuals are subject to random selection. All employees are included in the sample and are selected in 24 groups or clusters corresponding to the administrative units or sections into which the division is divided. Random times are selected from a list of three-digit random numbers. Random numbers from 000 through 509 are included to correspond to a work day beginning at 8:15 A.M. and ending at 4:45 P.M., 510 minutes later. Random numbers from 000 through 509 are added to 815 to obtain random times. For example, a random number of 369 corresponds to 8:15 plus 6 hours and 9 minutes or 2:24 P.M. These random times are assigned to the 24 units so that each of the latter is sampled at different times, although of course several of them may be sampled on the same day.

Random times are selected by the Sampling and Estimating Methods Section and turned over to the administrative office of the Division, and are telephoned to the supervisor or secretary of each unit as they occur. This is done as soon *after* the random time occurs as it is feasible to do so—usually within a minute or two. This call is a signal for every person in that unit to fill out a data record showing among other things what project he or she was working on at the time called. The supervisor or secretary is charged with the responsibility of seeing that a data record is filled out for every employee in the unit regardless of whether present or absent; and of reviewing data sheets to insure that they are free of gross errors.

THE DATA COLLECTING PROCEDURE

The data collecting procedure can best be explained in connection with a booklet giving instructions to employees, a list of projects by name and code number, and the data sheet, a copy of which is reproduced here. This coded list of projects and activities, and the data sheet which usually required only a few seconds to complete, combined with anonymous self reporting, constitute the major features of the data recording and collecting procedure.

Seven items of information are required on each data sheet. The first is the date and the second is the random time telephoned to the unit by the administrative office. The third is the salary grade, which can be checked on a list. The fourth is the project number, which is obtained from either coded list of projects or from the instruction booklet for employees if the activity is not covered by the project list. The fifth item is to determine whether the employee lacks a work assignment. The sixth item is to determine whether the

DISTRIBUTION OF COST SURVEY—DATA SHEET

Instructions: To determine the costs of the various projects engaged in by the Statistics Division, you will be asked periodically to complete the check sheet below showing how your time is being spent. DO NOT SIGN the sheet, but be as accurate as possible. This is not a measurement of your individual performance, but rather an attempt to determine more accurately the costs of the services rendered by the division.

1. DATE: _____

TIME: _____

2. CHECK (√) YOUR GRADE:

GS-1 ☐ GS- 9 ☐
GS-2 ☐ GS-10 ☐
GS-3 ☐ GS-11 ☐
GS-4 ☐ GS-12 ☐
GS-5 ☐ GS-13 ☐
GS-6 ☐ GS-14 ☐
GS-7 ☐ GS-15 ☐
GS-8 ☐

3. ENTER PROJECT NO. _____ **OR CHECK (√) IF NO WORK ASSIGNED** ☐.

4. CHECK (√) APPROPRIATE DESCRIPTIONS OF WHAT YOU WERE JUST DOING:

(1) ☐ Performing assigned subject matter duties (*Typing, coding, punching, supervising, writing, editing, planning, etc.*)
(2) ☐ Performing related duties (*Receiving instructions, laying out work, getting supplies, sharpening pencils, etc.*)
(3) ☐ Waiting (*For work, instructions, etc.*)
(4) ☐ Other than (1), (2) or (3)
(5) ☐ On leave
(6) ☐ At lunch

5. CHECK (√) YOUR SPECIFIC ORGANIZATION:

(1) ☐ Office of Director
(2) ☐ Administrative Office

PROGRAM BRANCH

(3) ☐ Office of Chief
(4) ☐ SOI Section
(5) ☐ Management Stat. Sec.
(6) ☐ Operation Reports Sec.
(7) ☐ SEM Section

OPERATIONS BRANCH

(8) ☐ Office of Chief
(9) ☐ Stat. Tech. Lab. Sec.

Presentations Section

(10) ☐ Office of Chief
(11) ☐ Graphics Unit
(12) ☐ Typing and Prf. Unit

Manual Operations Section

(13) ☐ Office of Chief
(14) ☐ Coding Unit
(15) ☐ Stat. Auditing Unit
(16) ☐ SOI Comp. and An. Unit
(17) ☐ Prod. Stat. Comp. and An. Unit
(18) ☐ Returns and Rec. Files Unit

Mechanical Operations Section

(19) ☐ Office of Chief
(20) ☐ Project Planning Unit
(21) ☐ Transcribing Unit
(22) ☐ Compiling Unit
(23) ☐ Punch Card File Unit
(24) ☐ Mach. Res. Rev. Unit

U. S. Treasury Department-Internal Revenue Service — Form M2079 (Rev. 6-20-55)

employee is in work status or not; if in work status, whether working on assigned duties or related duties, or waiting for work; and if not in work status, whether on leave or at lunch. The seventh item identifies to which of the 24 organizational units the employee is assigned.

The data sheets are verified to detect any gross errors such as wrong salary grade, wrong section, or an insufficient number of data records for a given section.

Does anonymous self reporting result in accurate data? Or conversely, does this method lead to distortion of the true work situation? Our experience covering a period of two years leads to the conclusion that if there is any falsification of the true work situation, it is so infrequent as to be negligible. The evidence for this statement is as follows:

1. Data records are anonymous.
2. No administrative use is made of individual data.
3. The random work sample has become routine since it has been carried on now for two years (November 1957).
4. Units which show certain high percentages, such as per cent time waiting for work, are those for which observation leads to the same conclusion.

THE METHOD OF ESTIMATION

The mechanics of deriving costs from the sample data required the development of a set of cost factors. These cost factors required the use of base pay before deductions for each grade level inasmuch as the budget with which the cost estimates were to be related was prepared on a similar basis. The assigning of time and cost values to each of the sample observations could then proceed on a straight line basis since the time period covered, the base salary paid during the period covered by each individual set of observations, and the number of observations made were known. By dividing the total time and cost by the number of sample observations a unit value of time and cost for each sample observation was obtained. This unit value is the same for every person with the same base salary and the same sampling rate.

For example, if the period of time covered was 5 days and the salary cost for 5 days was $100 and the number of sample observations during the 5 days was 3, then the values to be assigned to each sample observation was 5/3 days for time and $100/3 for cost. These values were introduced into a punched card which was prepared for each observation. These values were gang punched into each card for each person with the same base pay and the same sampling rate. By simple summation, the cost in time or money could then be obtained for any of the attributes appearing in the card.

ADVANTAGES OVER OTHER METHODS

The foregoing application of work sampling techniques includes a number of procedures which are an advantage over other methods.

1. The collection of salary data on the data record and the use of the method of estimation inherent in the sample design make it possible to estimate salary costs and their proper sampling errors. This procedure avoids the loss of information due to the use of a method of estimation divorced from the sample design.
2. The telephoning of random numbers, the use of anonymous self reporting, the use of an abbreviated data sheet, the use of codes for reporting activities, and the use of a simplified method of estimation all reduce drastically the cost of implementing the work sampling method without sacrificing the quality of the data obtained.

3. By making every minute subject to random sampling, and by including every person, it is possible to obtain estimates of time and cost which apply to the entire Division, and which therefore are applicable to budget planning and other problems involving the entire Division. Sampling some of the employees all of the time, or all the employees some of the time, or some of the employees some of the time, would not provide the information required.
4. A random time is called a minute or so after it actually occurs so that the sampling has no effect on what the person was doing. Furthermore, with the telphoning of random times, no one gets any advance cues, as is likely to be the case when a person makes a tour of an area in order to observe and record what employees are doing.
5. Continuous long range sampling means that variations between days within weeks, between weeks within months, and between months within a year, are given their proper weights in the estimates. This cannot be done where studies are limited to a few days, or even to a few weeks.
6. Sampling error formulas appropriate for cluster sampling are employed. Not all work sampling models are binomial requiring the use of PQ/n to estimate sampling variance.

EXAMPLES OF ESTIMATES AND THEIR SAMPLING ERRORS

The following examples illustrate methods of estimating proportions and aggregates and their sampling errors, using formulas described in the technical appendix. These examples are highly simplified for presentation purposes; in practice estimates would not be made from such small samples.

If the size of the random sample (number of clusters or number of individuals selected at random) is large enough, then we are warranted in assuming that the estimate (proportion or aggregate) is distributed approximately as the normal curve. Under these conditions we can calculate the confidence limits—estimate $\pm ks$—where k is the normal deviate corresponding to the confidence level (1.96 for 95 per cent level), and s is square root of the sampling variance of the estimate derived from the sample.

1. Assume the following array of activity A where 1 means presence of A and 0 its absence:

N	M (time instants)					Sum (time instants)
	1	2	3	4	5	
1	0	0	1	1	1	3
2	0	0	0	0	0	0
3	1	1	1	1	1	5
4	0	0	0	0	0	0
Sum Y_j	1	1	2	2	2	8
P_j	1/4	1/4	2/4	2/4	2/4	8/20

Since $M=5$ and $N=4$ the population proportion, equation (2) appendix, is

$$P = \tfrac{1}{5}(1/4 + 1/4 + 2/4 + 2/4) = .40.$$

Also

$$P = \frac{1}{NM}\sum_{1}^{N}\sum_{1}^{M} Y_{ij} = \frac{8}{4 \times 5} = .40.$$

Select two times at random, say the second and fourth time intervals, so that $m=2$. Then an estimate of $P=.40$, equation (3), appendix, is

$$p = \tfrac{1}{2}(1/4 + 2/4) = .375$$

also since $N=4$ and $m=2$

$$p = \frac{1}{Nm}\sum_{1}^{N}\sum_{1}^{m} y_{ij} = \frac{1+2}{4 \times 2} = \frac{3}{8} = .375.$$

2. The sampling error of the foregoing estimate of 0.375 is calculated as follows equation (5) appendix: where $M=5$, $m=2$, and $p=\frac{3}{8}$ or .375;

$$s^2 = \frac{(5-2)}{5} \times \frac{(1/4 - 3/8)^2 + (2/4 - 3/8)^2}{2(2-1)}$$

or

$$s = 0.096.$$

3. Estimate the proportion and its sampling error if time is sampled at random, and we center attention on a particular person. Assume $M=20\times480$ and $N=50$.

Let $m=20$ (once daily). Assume 16 data sheets show activity A.

Then

$$p = \frac{16}{20} = 0.80$$

and from equation 13 appendix:

$$s^2 = \frac{(20 \times 480 - 20) \times .80 \times .20}{20 \times 480 \times 19}$$

or

$$s = .09.$$

4. Assume random sampling of M only where $M=2000$ and $m=10$. The factor $(M-m)/M$ can be ignored. Assume that the 10 sample values of p_j are:

.09	.25
.15	.08
.12	.16
.20	.18
.11	.12

The average value

$$p = \frac{1.46}{10} = 0.146.$$

Then from equation 5 appendix the sampling error of this estimate is

$$s^2 = \frac{\sum_1^{10} (p_j - 0.146)^2}{10 \times 9}$$

or

$$s = 0.017.$$

5. Assume the following array of unit dollar costs for activity A for 8 persons. Let $M = 10 \times 480$ and $m = 10$.

N	m random times (minutes)										Sum
	1	2	3	4	5	6	7	8	9	10	
1					.02			.02	.02		.06
2	.02		.02	.02	.02		.02		.02		.12
3		.03				.03			.03		.09
4	.04					.04			.04		.12
5			.03				.03				.06
6				.02	.02			.02			.06
7											0
8											0
Total x_j	.06	.03	.05	.04	.06	.07	.05	.04	.11	0	.51

The estimated total cost is, equation (16) appendix,

$$\text{Total cost} = \frac{M}{m} \sum_1^m x_j = \frac{10 \times 480}{10} \times .51 = \$244.80.$$

The sampling error of this estimate, equation (19) appendix, where

$$\bar{x} = \frac{.51}{10} = .051,$$

is

$$s^2 = \frac{4800^2}{10 \times 9} \sum_1^{10} (x_j - .051)^2.$$

The value of s is \$34.64.

TECHNICAL APPENDIX

The activity array. Consider an activity array population of MN cells constructed from M time units and N persons, in which each cell is identified with some activity which may differ from unit to unit and from person to person.

THE ACTIVITY ARRAY

Individual	Time Units						
	1	2	3	...	j	...	M
1	A_{11}	A_{12}	A_{13}		A_{1j}		A_{1M}
2	A_{21}	A_{22}	A_{23}		A_{2j}		A_{2M}
3	A_{31}	A_{32}	A_{33}		A_{3j}		A_{3M}
.							
.							
.							
i	A_{i1}	A_{i2}	A_{i3}		A_{ij}		A_{iM}
.							
.							
.							
N	A_{N1}	A_{N2}	A_{N3}		A_{Nj}		A_{NM}

Let $y_{ij}=1$ if A_{ij} is the activity under study, and zero otherwise. Then the proportion of time spent on a particular activity in the jth cluster (time period) is

$$P_j = \frac{1}{N}\sum_{i=1}^{N} y_{ij}. \tag{1}$$

The proportion of time in the population spent on a particular activity over the total period of time is

$$P = \frac{1}{M}\sum_{1}^{M} P_j. \tag{2}$$

If a random sample of m clusters (time periods) is selected from a total of M clusters (time periods) and if all N are included, then an unbiased estimate of P the population proportion is

$$p = \frac{1}{m}\sum_{1}^{m} p_j \tag{3}$$

where p_j is the proportion in the jth sampled cluster.

The true variance, $V(p)$, of the estimate p is given by the following expression:

$$V(p) = \frac{M-m}{M}\,\frac{\sum_{J=1}^{M}(P_j - P)^2}{m(M-1)}. \tag{4}$$

The unbiased estimate of $V(p)$, is given by $v(p)$ where

$$v(p) = \frac{M - m}{M} \frac{\sum_{1}^{m} (p_j - p)^2}{m(m-1)}. \tag{5}$$

The factor $(M-m)/M$ can be ignored if m is very small (less than 5%) compared with M.

For stratified random sampling the proportion of time spent on a particular activity over the whole period of time for all k strata in the population is as follows:

$$P_{st} = \frac{1}{N} \sum_{j=1}^{k} N_j P_j' \tag{6}$$

where N_j is the number of elements in each cluster in the jth stratum, $N = \sum_1^k N_j$, and P_j' is the proportion of time spent on a particular activity over the total period of time in the jth stratum.

An unbiased estimate of P_{st} is

$$p_{st} = \frac{1}{N} \sum_{1}^{k} N_j p_j' \tag{7}$$

where p_j' is an unbiased estimate of P_j', and is obtainable from (3).

The true variance,

$$V(p_{st}) = \frac{\sum_{j=1}^{k} N_j^2 V(p_j')}{N^2} \tag{8}$$

where $V(p_j')$ is the true variance of the proportion in the jth stratum, and is given by (4).

An unbiased estimate of $V(p_{st})$ is given by $v(p_{st})$ where

$$v(p_{st}) = \frac{\sum_{j=1}^{k} N_j^2 v(p_j')}{N^2} \tag{9}$$

and $v(p_j')$ can be obtained from equation (5) above.

If estimates of the proportion of time spent by an *individual* on a particular activity are also of interest, then the following formulas are appropriate. Note that we are still sampling clusters; however in this case since interest is centered in an individual, each cluster will have but two values, 1 or 0, depending on whether individual i, say, is working on a particular activity or not.

The proportion of time individual i spends on a particular activity over the total period of time is

$$P_i = \frac{1}{M} \sum_{j=1}^{M} Y_{ij}. \tag{10}$$

An unbiased estimate of P_i is p_i where

$$p_i = \frac{1}{m} \sum_{j=1}^{m} y_{ij} \tag{11}$$

and y_{ij} is the sample value of Y_{ij}.

The true variance $V(p_i)$ of the estimate p_i is

$$V(p_i) = \frac{M - m}{M - 1} \frac{P_i Q_i}{m} \text{ where } Q_i = 1 - P_i. \tag{12}$$

An unbiased estimate of $V(p_i)$ is given by $v(p_i)$ where

$$v(p_i) = \frac{M - m}{M} \frac{p_i q_i}{m - 1} \text{ and } q_i = 1 - p_i. \tag{13}$$

The unit cost array. A logical counterpart of the activity array is the unit cost array. In this array there are MN cells the same as in the activity array, but each cell entry is the salary cost (C_{ij}) assignable to a particular person and a specified time unit.

THE UNIT COST ARRAY

Individual	Time Units						
	1	2	3	...	j	...	M
1	C_{11}	C_{12}	C_{13}		C_{1j}		C_{1M}
2	C_{21}	C_{22}	C_{23}		C_{2j}		C_{2M}
3	C_{31}	C_{32}	C_{33}		C_{3j}		C_{3M}
.							
.							
.							
i	C_{i1}	C_{i2}	C_{i3}		C_{ij}		C_{iM}
.							
.							
.							
N	C_{N1}	C_{N2}	C_{N3}		C_{Nj}		C_{NM}

Let C_{ij} represent the cost per unit of time associated with individual i at time j. Let $X_{ij} = C_{ij}$ if activity A_{ij} is under consideration, but equal to zero otherwise.

Then

$$X_{ij} = \sum_{i=1}^{N} X_{ij} \tag{14}$$

gives the aggregate salary cost in the jth cluster (column) for a particular activity. Also

$$X = \sum_{j=1}^{M} X_j \tag{15}$$

gives the aggregate salary cost for the particular activity for the total period of time.

If a random sample of m clusters is selected from a total of M clusters and if all N are included, then

$$x' = \frac{M}{m} \sum_{j=1}^{m} x_j = M\bar{x} \tag{16}$$

is an unbiased estimate of the population aggregate X, where x_j is the sample value of X_j and where

$$\bar{x} = \frac{\sum_{1}^{m} x_j}{m}.$$

The true variance, $V(x')$, of the estimated aggregate x' is

$$V(x') = \frac{M(M-m)}{m(M-1)} \sum_{j=1}^{M} (X_j - \overline{X})^2 \tag{17}$$

where $\overline{X}$ is

$$\frac{\sum_{1}^{M} X_j}{M}.$$

An unbiased estimate of $V(x')$ is given by $v(x')$ where

$$v(x') = \frac{M(M-m)}{m(m-1)} \sum_{j=1}^{m} (x_j - \bar{x})^2. \tag{18}$$

If the sample m is small compared with M then this expression becomes:

$$v(x') = \frac{M^2}{m(m-1)} \sum_{j=1}^{m} (x_j - \bar{x})^2. \tag{19}$$

If stratified random sampling is used we apply equation (16) to each stratum and add the k values of x' in order to obtain the estimated total aggregate. Similarly we apply equation (18) to every stratum and add the k values of $v(x')$ in order to obtain the total sampling variance of the estimated total aggregate.

REFERENCES

[1] Abruzzi, A., *Work Measurement.* New York: Columbia University Press, 1952, 246–72.

[2] Barnes, Ralph M., *Work Sampling.* Dubuque, Iowa: Wm. C. Brown Company, 1956.

[3] Brisley, C. L., "Work sampling (ratio-delay study) technique," *Quality Control Conference Papers* (1953), American Society for Quality Control, 115–25. See also Brisley, "How you can put work sampling to work," *Factory Management and Maintenance,* July 1952, 84–9.

[4] Cochran, W. G., *Sampling Techniques.* New York: John Wiley & Sons, 1953.

[5] Morrow, R. L., *Time Study and Motion Economy.* New York: The Ronald Press, 1946, 175–99. See also *Mechanical Engineering,* April 1941, 302–3.

[6] Rand Corporation, *A Million Random Digits with 100,000 Normal Deviates.* Glencoe, Illinois: The Free Press, 1955.

[7] Tippett, L. H. C., "A snap-reading method of making time-studies of machines and operators in factory surveys," *The Journal of the Textile Institute Transactions,* February 1935, 51–70.

THE TRENTILE DEVIATION METHOD OF WEATHER FORECAST EVALUATION*

MORRIS JAMES SLONIM
Headquarters, United States Air Force

This article outlines a method originated by the author during World War II for evaluating short range weather forecasts. An interesting feature is the use of "deviational penalty scores" to adjust for the relative ease or difficulty of forecasting for varying weather situations. A number of other studies on weather forecast evaluation by other authors lead to different methods of analysis (see references at end of article). The author has encountered situations in other fields where variations of the method proved useful. As an instance, some elements of this method were used in designing a scoring procedure for rating bombing proficiency of air crews during the Korean War.

DEVELOPMENT OF THE SCORING METHOD

IN order to evaluate the accuracy of weather forecasts, it is necessary, first, to take account of the degree of difference between forecast and observed values. Thus, generally speaking, a person forecasting atmospheric pressure within one millibar of the observed value should get a better score for that particular forecast than a person whose prediction was off by ten millibars. In addition to taking account of the degree of difference, however, it is essential that account be taken of the relative difficulty or ease of forecasting various values of a given meteorological element. For example, it is much easier to forecast correctly "no rain" for the six-hour period preceding 1230 Z-time[1] during April for Jacksonville, Florida, than it would be to forecast correctly .01 inch of rain for the same place for the same time. The point is illustrated by the following tabulation which shows the actual distribution of precipitation in Jacksonville for the six-hour period 0630 Z to 1230 Z for April over a six-year period:

TABLE 398

Precipitation Amounts (in inches)	Number of Observations
No Rain	147
Trace	15
.01	6
.02–.15	6
.16 or more	6
Total	180

It can be observed from the above sample distribution that a person forecasting "no rain" has a chance of 147 in 180 of predicting correctly; whereas

* The author is particularly indebted to Lewis T. Reinwald of the Corporation for Economic & Industrial Research for suggestions made for this article.

[1] Z-time stands for Greenwich Mean Time.

the chance that a prediction of .01 inch of rain will be correct is only 6 in 180. It would seem logical that two persons correctly forecasting these two precipitation values from different six-hourly synoptic maps should not be given the same score.

In order to take into account the degree of difference between forecast and observed, and at the same time to take account of the relative difficulty or ease of forecasting any meteorological element in different parts of the range, the following approach was taken.

Denote the observed values of a specified variable meteorological element (e.g., precipitation, temperature, atmospheric pressure) by X with probability density function $f(X)$. Let Y be the predicted value corresponding to each observed X. If we divide the frequency distribution of X into m intervals so that $f(X)$ for each interval is $1/m$ we have transformed our original distribution to a uniform distribution. In the simplest situation we will have different values of X in each of the m equal probability intervals. In such a situation, these intervals $I_1, I_2, I_3, \cdots, I_m$ can be assigned the successive location values of 1, 2, 3, $\cdots$, m. Then, any given forecast (Y) as well as the actual observed value (X) can be assigned the corresponding location values m_X and m_Y and the score $S(X, Y) = |m_X - m_Y|$. Each distribution was divided into 30 equal probability intervals (or trentiles).[2] The number 30 was selected rather arbitrarily, and represents an attempt to secure a spread in scores while keeping the number of intervals small enough to be manageable for machine processing. It developed also that 30 intervals was about the maximum for a legible forecast form (See Figure 398).

This method of scoring automatically takes account of the degree of difference between forecast and observed values and thereby assigns a perfect score (zero) to a "hit" and a score of 1, 2, $\cdots$, $m-1$ to a "miss," with the highest scores assigned to the widest "misses." More particularly, the scoring method takes account of the relative ease or difficulty of predicting a specific value accurately by converting the climatological data for a specified element into intervals of equal probability.[3]

Following is an example of a table of trentiles for a distribution of atmospheric pressure values.

The individual forecaster's score on his prediction of atmospheric pressure for 0030 Z time for the week of April 5–April 12 for Seattle would be obtained by converting the difference between forecast and observed values into trentiles in the table and getting their difference. Let us assume a forecaster predicted a pressure of 1011.8 millibars for 0030 Z for Seattle for April 7th. From the table it can be seen that if the observed value let us say, turned out to be 1017.6 millibars, the forecaster's score for this portion of his forecast would be 10 (i.e., 15−5).

---

[1] For the sake of euphony and brevity the author took the liberty of using the French rather than the more conventional Latin prefix in forming the word "trentile."

[2] Frequency distributions of the values of each of the five elements forecast were obtained from observations recorded for previous years for each station at each of the Z-times for the 30-day period centering about the week in question. The five elements selected for the verification system were atmospheric pressure, temperature, ceiling height, visibility, and precipitation.

TABLE 400

Trentile Number	Atmospheric Pressure Values (in Millibars)
1	less than & including 1004.4
2	1004.5 to 1007.8
3	1007.9 to 1009.5
4	1009.6 to 1010.5
5	1010.6 to 1012.2
6	1012.3 to 1013.2
7	1013.3 to 1013.5
8	1013.6 to 1014.2
9	1014.3 to 1014.9
10	1015.0 to 1015.2
11	1015.3 to 1015.6
12	1015.7 to 1015.9
13	1016.0 to 1016.9
14	1017.0 to 1017.3
15	1017.4 to 1018.0
16	1018.1 to 1018.3
17	1018.4 to 1018.6
18	1018.7 to 1019.0
19	1019.1 to 1019.7
20	1019.8 to 1020.3
21	1020.4 to 1021.0
22	1021.1 to 1021.3
23	1021.4 to 1022.4
24	1022.5 to 1023.4
25	1023.5 to 1024.0
26	1024.1 to 1024.4
27	1024.5 to 1025.4
28	1025.5 to 1027.1
29	1027.2 to 1029.1
30	1029.2 and over

The above trentile distribution for pressure was a simple example, purposely selected to illustrate the general method employed. It will be noticed that no single pressure value falls in more than one of the trentiles. To illustrate the procedure that should be followed when a single value of an element, because of its frequency of occurrence, occupies more than one trentile, the following table of trentiles was selected.

It is apparent that a method is needed that will yield different trentile deviations between predicted and observed for correct forecasts of values occupying more than one trentile (such as "no rain" or "trace" in Table 401, above) than it will yield for correct forecasts of values that are included in only one trentile (such as .01 inch). Thus, for a correct forecast of .01 inch of rain (24th trentile), the deviational score is zero (24 minus 24). If, however, a person forecasts "no rain" and "no rain" is observed, it would obviously be unfair to give him a perfect score of zero, inasmuch as the likelihood of making a correct random forecast of .01 inch was only 1 in 30, as against 20 in 30 in the case of pre-

TABLE 401

Trentile Number	Amounts of Precipitation (in inches)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20	No Rain
21 22 23	Trace
24	.01
25	.02
26	.03–.04
27	.05–.08
28	.09–.13
29	.14–.21
30	.22 and over

dicting "no rain" correctly. The method used to differentiate between two such correct forecasts in assigning scores was termed the method of "deviational penalties" and is based upon the following reasoning:

Assume, for purposes of illustration, that precipitation forecasts and observations can be measured within microscopic limits so that the value "no rain" can be broken into as many minute class intervals as there are trentiles in the "no rain" category. If "no rain" occupies 2 trentiles the following possibilities exist:

Forecast Value May Fall in Trentile:	Observed Value May Fall in Trentile:	Trentile Deviation
1	1	0
1	2	1
2	1	1
2	2	0
Total Deviation		2

The average deviation is $\frac{1}{2}$: i.e.,

$$\frac{\text{Sum of Deviations}}{\text{No. of Deviations}} = \frac{2}{4}$$

The formula for the general case, where the number of trentiles embracing a single value of the element is N can be derived in the following way:

D = Deviational Penalty
N = Number of trentiles occupied by a single value of meteorological element.
n_i = Predicted trentile
n_j = Observed trentile
N^2 = Total number of possible different pairs of forecast-and-observed values.

$$D = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} |n_i - n_j|}{N^2} \tag{1}$$

But

$$\sum_{i=1}^{N}\sum_{j=1}^{N} |n_i - n_j| = 2\sum_{i=1}^{N}\sum_{j=1}^{n_i} (n_i - n_j) \tag{2}$$

since

$$|N-1| + |N-2| + \cdots + |2-N| + \cdots + |1-N| + \cdots$$
$$= (N-1) + (N-2) + \cdots + (-1)(2-N) + \cdots + (-1)(1-N) + \cdots$$
$$= 2(N-1) + 2(N-2) + \cdots$$

Moreover,

$$2\sum_{i=1}^{N}\sum_{j=1}^{n_i} (n_i - n_j) = 2\sum_{i=1}^{N}\left[n_i^2 - \frac{n_i(n_i+1)}{2}\right] \tag{3}$$

since

$$\sum_{j=1}^{n_i} n_i = n_i \sum_{j=1}^{n_i} 1 = n_i^2$$

and

$$\sum_{j=1}^{n_i} n_j = \frac{n_i(n_i+1)}{2}$$

Simplifying, (3) becomes

$$2\sum_{i=1}^{N}\left(\frac{n_i^2}{2}-\frac{n_i}{2}\right)=\frac{N(N+1)(2N+1)}{6}-\frac{3N(N+1)}{6}=\frac{N(N^2-1)}{3}$$

Hence

$$D=\frac{N(N^2-1)}{3N^2}=\frac{N^2-1}{3N} \tag{4}$$

This value, $(N^2-1)/3N$ is the "deviational penalty" trentile score assigned to a forecaster who correctly predicts a value of an element that occupies more than one trentile. The penalty score is based on the reasoning that if forecasts and observations were made within extremely fine limits the forecaster's average score over all possible different random forecast-and-observed situations within a group of trentiles including a single value would be $(N^2-1)/3N$. In the precipitation distribution of Table 401 above, for example, a person forecasting "no rain" would, if "no rain" were observed, get a score of 6.6, since "no rain" occupies 20 trentiles (i.e. $N=20$).

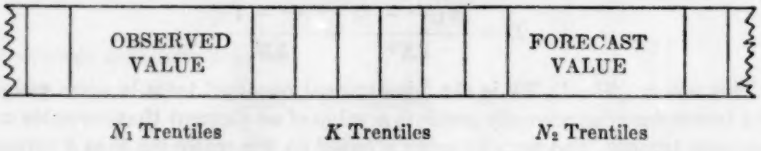
Another situation that occurs frequently and which requires separate scoring treatment is one in which the forecast value differs from the observed value and either the forecast or observed values, or both, occupy more than one trentile. Let us examine, for example, the following distribution of temperature:

TABLE 403

Trentile No.	Temp. in °F
1 2 3 4 5	43°
6	44°
7	45°
8	46°
9	47°
10	48°
11 12 13	49°
etc.	etc.

The formula for this type of scoring situation can be derived as follows:
Let

T = Trentile score
N_1 = Number of trentiles occupies by the observed value
n_i = Number of ith observed trentile
N_2 = Number of trentiles occupied by the forecast value
n_j = jth forecast trentile
K = Number of trentiles between forecast and observed ranges
N_1N_2 = Total number of possible different pairs of forecast-and-oberved



$$T = \frac{\sum_{j=1}^{N_2}\sum_{i=1}^{N_1}(N_1 + K + n_j) - n_i}{N_1N_2} \tag{1}$$

$$= \sum_{j=1}^{N_1}\left[N_1(N_1 + K + n_j) - \frac{N_1(N_1+1)}{2}\right] \Big/ N_1N_2$$

$$= \left[N_1N_2(N_1 + K) + \frac{N_1N_2(N_2+1)}{2} - \frac{N_1N_2(N_1+1)}{2}\right] \Big/ N_1N_2 \tag{2}$$

$$= \frac{N_1 + N_2}{2} + K$$

(The forecast and observed trentile ranges may of course be contiguous in which case K equals zero).

Referring to Table 403 above, if a temperature forecast is 49°F and the observed is 43°F, the trentile score is 9, since $N_1=5$, $N_2=3$, and $K=5$.

The use of simple differences in trentile values when both forecast and observed fall into single trentiles; the use of $(N^2-1)/3N$ when a correct forecast value occupies more than one trentile; and the application of

$$\left(\frac{N_1 + N_2}{2}\right) + K$$

for incorrect forecasts where forecast or predicted values or both occupy more than one trentile, together permit the scoring of all possible combinations of forecasts and observations.

The trentile deviation score was intended as a skill score and the scoring system was designed so that the lowest trentile deviation scores would reflect the highest levels of forecasting skill. The "perfect" score for a series of completely accurate forecasts would be zero if all observed values occupied single trentiles. Otherwise, the "perfect" score would be the sum of the deviational penalties. This would represent the minimum score and optimum performance for any forecast.

Forecasting skill as reflected by the normalized trentile deviation scores[4] may be gauged by comparing a forecaster's average score for a number of forecasts with the average score obtained by the use of methods requiring no specialized training in meteorology. A forecaster could be said to possess some degree of skill if his average score were significantly higher (for the standardized trentile deviation scores) than the highest of the "unskilled" forecast average scores. Such "unskilled" scores were actually calculated daily during most of the operation of the "Trentile Deviation" program. Scores, for example were computed for "chance" forecasts by selecting one of the 30 trentiles at random for each of the 25 forecast situations daily; for "climatological" forecasts, which were prepared by predicting the maximum frequency value for each of the five elements for each of the five cities each day; and for "persistence" forecasts by predicting that the weather element values would remain the same over the period of the forecast.

THE SHORT RANGE WEATHER FORECAST VERIFICATION FORECAST FORM

When the Short Range Verification Program was inaugurated, in April, 1943, the forecasters made 24-, 36-, and 48-hour forecasts for five selected stations (which were changed weekly), for atmospheric pressure, temperature, ceiling height, visibility, and precipitation, using blank forecast forms.

After the program had operated for 20 weeks, a revised program was inaugurated in October, 1943. A new short range forecast form, Figure 407, was devised. The new form supplied the forecaster with a certain amount of climatological background showing the climatological distribution of the elements to be forecast. This form also simplified the preparation and processing of the forecasts, so that virtually the entire scoring procedure was now performed rapidly and economically by punch card machine methods.

In addition, the requirements were changed (when the new program was initiated) to provide for the use of only one synoptic map daily instead of four and to provide for forecasts covering only one period weekly instead of three. The period of forecast was changed each week to include 24-, 48-, 12-, and 36-hour forecasts, in that order, in each four-week interval. The five stations were changed every four weeks. The same five meteorological elements were retained for the duration of the project.

SCORING THE FORECASTS

Each forecaster was requested to complete a minimum of three forecast forms each week. The forecast form (Figure 398) provides for forecasting five elements for each of five stations—a total of 25 separate predictions. Each of the 25 predictions was scored according to the "trentile deviation" system described above, and the forecaster's total "trentile deviation" score for each form was obtained by summing these 25 individual forecasts. Since the scores for all elements and for all the stations were in terms of a uniform measure, the trentile scores for the separate stations and for the separate elements could properly be combined into a single total score.

---

[4] As noted in a later paragraph, the trentile deviation scores were subsequently converted to standard scores in order to equalize the effect of "easy" and "difficult" forecast days.

Since all forecasters in the continental United States forecast for the same five elements, for the same five stations, from the same synoptic maps (1230 Z), their scores in any given day were believed to be comparable.

In the early weeks of the program each forecaster's weekly (or "to date") average score for a given length of forecast was obtained by adding his "trentile deviation" scores for the period in question and obtaining his average score per forecast form. Analysis of the daily distributions of "trentile deviation" scores indicated that the scoring system did not automatically adjust sufficiently for the variations in synoptic situations from day to day; so that some forecasters might be unduly rewarded while others might be overly penalized in the average rankings as a result of the daily changes in the relative ease or difficulty of forecasting from the varying synoptic situations. (See reference [14] for a review of the results of the first six months of the verification program.)

Accordingly, all "trentile deviation" scores for each day were converted to standard scores. In this way the scores for each day were made comparable with those of the other days in which forecasts of the same duration were made. Some 500,000 short range weather forecasts were evaluated from inauguration of the program until the end of World War II using the "trentile deviation" system.

REFERENCES

[1] Clayton, H. H., "A method of verifying weather forecasts," American Meteorological Society, *Bulletin* 8 (1927), pp. 144–6.

[2] Clayton, H. H., "Rating weather forecasts," American Meteorological Society, *Bulletin* 15 (1934), pp. 279–83.

[3] Doolittle, M. H., "The verification of predictions," Philosophical Society of Washington, *Bulletin* 7 (1885) pp. 122–7.

[4] Ekholm, Nils, "Über die unperiodischen Luftdruckschankungen and einige damit zusammenhangende Erscheinungen (VI-Die Wetterprognose)," *Meteorologische Zeitschrift* 24 (1907), pp. 145–59.

[5] Gringorten, I. I., "The verification and scoring of weather forecasts," *Journal of the American Statistical Association*, Vol. 46 (1951), pp. 279–96.

[6] Heidke, P., "Berechnung des Erfolges und der Gute der Windstarkevorhersagen im Sturmwarnungsdienst," *Geografika Annaler* 8 (1926), pp. 310–349.

[7] Koppen, W., "Eine Neue Methode der Prufung der Wetterprognosen," *Meteorologische Zeitschriff* 1 (1884), pp. 39–40.

[8] Mascart, Jean, "La Proportion des reussites dans la prevision du temps," Academic des Sciences, *Comptes Rendus* 175 (1922), pp. 627–9.

[9] Milham, W. I., *Meteorology* (New York: Macmillan, 1912), pp. 404–8.

[10] Muller, R. H., "Verification of short range weather forecasts (a survey of the literature)," *Bulletin of the American Meteorological Society*, Vol. 25 (1944), pp. 18–25, 47–53, 88–95. (This comprehensive bibliography is especially valuable because of the excellent summaries of listed articles).

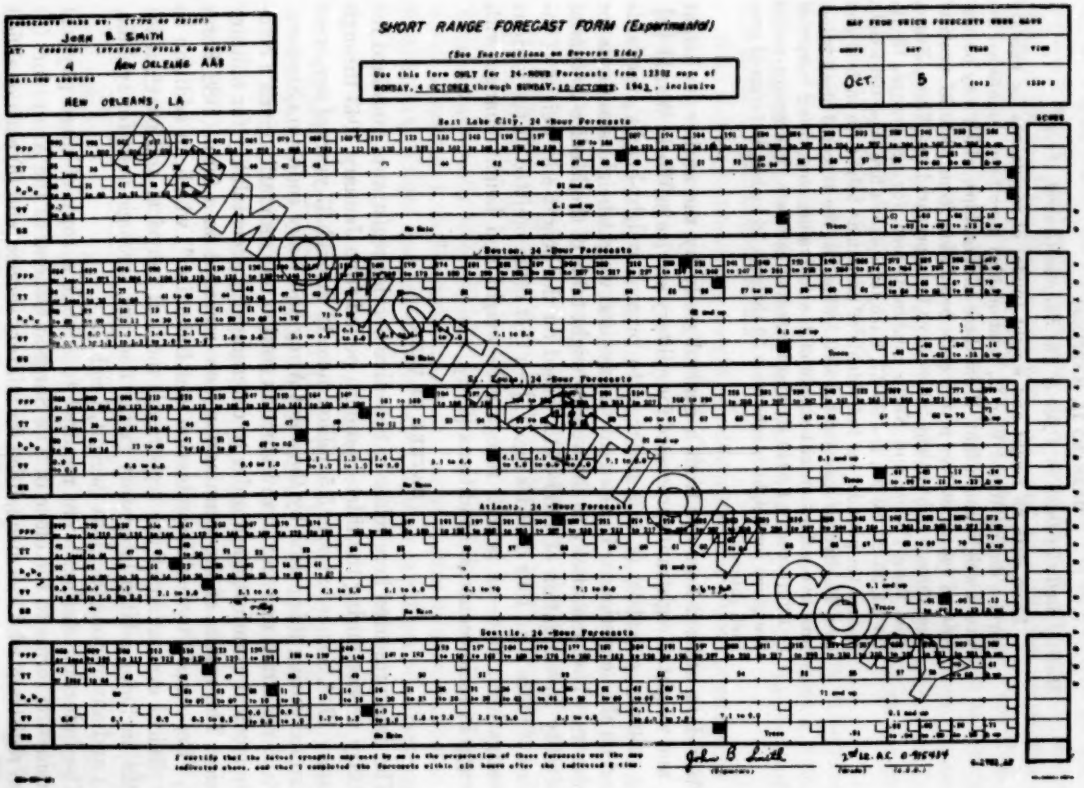
[11] Shaw, Sir Napier, "Forecasting weather," 3d Ed., (London, 1940). Chapter XXV, pp. 546–50.

[12] Slonim, M. J., "Method of Rating Short Range Forecasts" June 7, 1943 (Dittoed report by Weather Information Branch, Hq., Army Air Forces).

[13] U. S. Army Air Forces, Headquarters, Weather Information Branch, "Short Range Forecast Verification Program," Report No. 602, November 1943.

[14] U. S. Army Air Forces, Headquarters, Weather Information Branch, "Short Range Forecast Verification Program," Supplement I to Report No. 602, June 1944.

[15] Van Everdingen, E., "De Methode van Beoordeeling der weersver Wachtingen van het K. Ned. Met. Inst., en de Uitkomsten in den Zomer Van 1911," Netherlands Met. Inst., *Mededeelingen en Verhandelingen* 14 (1912).


FORECASTS MADE BY: (TYPE OR PRINT)
JOHN B. SMITH
AT: (AIRPORT, STATION, FIELD OR BLDG)
4 NEW ORLEANS AAS
MAILING ADDRESS
NEW ORLEANS, LA
SHORT RANGE FORECAST FORM (Experimental)
(See Instructions on Reverse Side)
Use this form ONLY for 24-HOUR Forecasts from 1230Z maps of MONDAY, 4 OCTOBER through SUNDAY, 10 OCTOBER, 1943, inclusive
MAP FROM WHICH FORECASTS WERE MADE
MONTH
DAY
YEAR
TIME
Oct.
5
1230 Z
Salt Lake City, 24 -Hour Forecasts
Boston, 24 -Hour Forecasts
St. Louis, 24 -Hour Forecasts
Atlanta, 24 -Hour Forecasts
Seattle, 24 -Hour Forecasts
SCORE
PPP
TT
Td-Tc
VV
RR
No Rain
Trace
I certify that the latest synoptic map used by me in the preparation of these forecasts was the one indicated above, and that I completed the forecasts within six hours after the indicated G.C.T.
John B. Smith
(Signature)
2nd Lt. A.C. 0-426414
(Rank)
(A.A.F.)

DEMONSTRATION COPY

Fig. 407

WEIGHT-HEIGHT STANDARDS BASED ON WORLD WAR II EXPERIENCE

BERNARD D. KARPINOS

Medical Statistics Division, Office of The Surgeon General, Army

From data on height and weight (nude) of registrants for military service during World War II, relationships have been established between weight and height. From these relationships, which may be considered representative of the male population from 18 to 37 years of age, three weight-height tables have been prepared for White, Negro, and Total (White and Negro combined) men, for age groups, 18–19, 20–24, 25–29, 30–34, and 35–57. Each table shows five expected-weight values (in pounds) for individuals of particular height (in inches) and age. The five values consist of a midvalue, i.e., the expected mean value, and two values on either side of the midvalue, comprising ranges of ± 1.2816 and ± 1.6449 standard errors of estimate. The weight-values are given for height-ranges from 59 through 78 inches, in terms of one-inch intervals.

THE data utilized here are from a general study by the author on height and weight of registrants processed for military service in World War II [4]. The primary objective of this paper is to present weight-height standards based on that experience. The current widely-used weight-height standards for adults trace back essentially to the weight-height table of the Medico-Actuarial Mortality Investigation of 1909–1912, based on the height and weight of insurees of more than four decades ago [1]—a fact often omitted in the various investigations. But—as is well recognized—weight and height change with changes in socio-economic conditions [2, 8].

NATURE OF DATA

The data presented here relate to Selective Service registrants examined for military service during the thirteen-month period from January 1943 through January 1944. Somewhat over 5.5 million men, 18 through 37 years of age, were examined during this period by the Joint Army and Navy Induction Stations.

Each registrant arriving at an induction station was ordinarily given a complete medical examination in order to determine his acceptability for military service. The results of the examination were recorded on a form (DSS Form 221, "Report of Physical Examination and Induction") which, in addition to the clinical evaluation of the registrant's medical (physical and psychiatric) status, included related laboratory (x-ray findings, serologic tests for syphilis, etc.) and anthropometric data (weight, height).

The induction stations were required to submit a copy of these physical examination forms to the office of The Surgeon General of the Army, for: (*a*) Each registrant who was disqualified for military service, and (*b*) each qualified registrant to be inducted into the Army. During this period about three-fourths of the qualified registrants called for induction were assigned to the Army and about one-fourth to the Navy. Copies of the physical examination forms for Navy inductees were not available to The Surgeon General's Office. Adjust-

ments, explained later, were therefore made for this partial coverage of the inducted registrants in order to reflect the total examined registrant population.

Each physical examination form carried a Selective Service number assigned to the registrant by his local board. From all the forms received during this period a ten-per cent random sample was drawn by choosing a single terminal digit of the Selective Service numbers and using all forms for which this was the terminal digit. About 465,000 medical examination forms were thus obtained, relating to individuals of known age and of known height and weight. A distribution of these forms by age, race, and military qualification (inducted or disqualified) of the individuals included in the analysis is shown in Table 409. Age is given as of last birthday. The term "White" used in the racial classification refers to Non-Negro.

The height and weight data for this study were taken from these sample forms. The examinees were measured without shoes and weighed without clothing. Height was recorded in inches and weight in pounds, as specified on the physical examination forms.

With respect to measuring height, specific instructions were provided by the Army regulations on physical standards (Mobilization Regulations 1–9, 1942) which read as follows:

> "Use a board at least 2 inches wide by 80 inches long, placed vertically, and carefully graduated to ¼ inch between 58 inches from the floor and the top end. Obtain the height by placing vertically, in firm contact with the top of the head, against the measuring rod an accurately squared board of about 6 by 6 by 2 inches, best permanently attached to graduated board by a long cord. The individual should stand erect with back to the graduated board, eyes straight to the front." [11]

Height was usually recorded in whole inches, though it was sometimes given to a half or quarter inch. Recorded fractions of an inch were dropped in coding the height data. Weight was presumably taken to the nearest pound. Fractions of

TABLE 409

NUMBER OF PHYSICAL EXAMINATION FORMS INCLUDED IN THE SAMPLE STUDY OF HEIGHT AND WEIGHT, BY AGE, RACE, AND MILITARY QUALIFICATION OF THE EXAMINEES

(January 1943 through January 1944)*

Age	White		Negro	
	Inducted	Disqualified	Inducted	Disqualified
18–19	77,744	32,507	9,171	6,968
20–24	67,824	29,971	11,559	9,572
25–29	40,934	28,971	6,661	9,446
30–34	34,567	34,225	5,407	10,897
35–37	16,303	22,891	2,535	6,422
Total (18–37)	237,372	148,565	35,333	43,305

* Source: Sample tabulations of DSS Form 221 ("Report of Physical Examination and Induction").

pounds, when recorded on the forms were disregarded, and weight was coded in whole numbers as in the case of height.

WEIGHT-HEIGHT RELATIONSHIP

In order to establish the relationship between weight and height, cross tabulations of weight by height were made for each subgroup, as shown in Table 409. Weight was expressed in the tabulations in 5-pound intervals, against height in one-inch intervals.

Based on these bivariate distributions, the means, standard deviations, and covariances, required for correlation analysis, were computed for each subgroup. The logarithms of weight, rather than the actual weights, and the actual heights were used in these calculations. The choice of this procedure was prompted by the following considerations. First, the F test for linearity of regression (Kendall [5], p. 234) applied to some of the subgroups, by linearly correlating actual weights and actual heights, revealed significant departures from linearity. (A similar phenomenon was found by Rosenbaum [10] in his weight-height investigation of English recruits.) Second, the standard deviations of weight for each inch of height tended to increase with height for each subgroup, while the corresponding coefficients of variation proved to be fairly constant, thus suggesting semilogarithmic regressions. (See Kenney and Keeping [7], p. 210; O'Brien, Girshick, and Hunt [9], p. 70, state that an analysis of their data indicated a semilogarithmic relationship of weight to height.)

Differences were found in the height and weight values of the various subgroups, whether differentiated by age, or race, or military qualification (inducted and disqualified).

ADJUSTMENTS FOR DISPROPORTIONALITY

As previously indicated, copies of the physical examination forms of the Navy inductees were not available to The Surgeon General's Office. Since differences were found between the weight-height values of qualified registrants (inductees) and those of disqualified registrants, direct additions of the bivariate distributions of these subgroups—for the purpose of obtaining corresponding distributions for the examined (inducted plus disqualified) groups—would have given undue weight to the disqualified subgroups and would have introduced thus a certain bias into the calculations. (For instance, for the (18–19) white age-group, the ratio of inducted to disqualified, as tabulated, was 71:29, versus an "expected" ratio of 76:24 for qualified versus disqualified; See Appendix, Table 417.)

To correct for these disproportionalities, the various values: means, standard deviations, and covariances derived from the bivariate distributions of the inducted and disqualified subgroups were weighted by their respective "expected" ratios (Table 417). (See "Appendix: Technical Note" for the rationale and method of computation.) The assumption was made in this case that all qualified registrants were distributed by weight and height as the Army inductees, since neither height nor weight nor other related factors played a decisive part in the assignment of qualified registrants to the Army or the Navy.

TABLE 411

COEFFICIENTS OF CORRELATION, REGRESSION EQUATIONS AND STANDARD ERRORS OF ESTIMATES OF WEIGHT (W) ON HEIGHT (H) OF REGISTRANTS PROCESSED FOR MILITARY SERVICE DURING WORLD WAR II, BY AGE AND RACE

(January 1943 through January 1944)*

Age and Race	Correlations	Regression Equations: $Log_{10}W = a + bH$	Standard Error of Estimate (Log.)
White			
18–19	.4770	1.38468+0.01129H	0.05330
20–24	.4336	1.45443+0.01048H	0.05579
25–29	.3954	1.48355+0.01029H	0.06126
30–34	.3780	1.50883+0.01003H	0.06363
35–37	.3810	1.49814+0.01025H	0.06434
Negro			
18–19	.5386	1.39879+0.01108H	0.04387
20–24	.5320	1.43190+0.01082H	0.04455
25–29	.4974	1.42707+0.01105H	0.04968
30–34	.4651	1.43843+0.01098H	0.05448
35–37	.4408	1.45003+0.01088H	0.05774
Total			
18–19	.4826	1.38623+0.01126H	0.05239
20–24	.4419	1.45210+0.01052H	0.05477
25–29	.4039	1.47730+0.01038H	0.06027
30–34	.3855	1.50139+0.01013H	0.06284
35–37	.3860	1.49341+0.01031H	0.06379

* Coefficients of correlation and regression equations were derived by correlating the logarithms of weight (Log_{10} W—in pounds) with height (H—in inches).

The correlation coefficients, regression equations, and standard errors of estimate, given in Table 411 by age and race (white, Negro), were derived from these adjusted values.

With these adjustments, the derived weight and height values, by age and race (Table 411), may be regarded as representative not only of the examined registrants, but of all registrants within the age range, though they exclude registrants disqualified by the local boards prior to the induction processing, as well as deferred and exempt registrants. During this period the local boards played a minor role in the screening of registrants for military service, so that the exclusion of registrants disqualified by the local board could have not produced an appreciable distorting effect on the weight-height data. As to the deferred and exempt registrants, there is no reason to believe that their distributions by weight and height differed materially from that of the other registrants. There are indications that such differences would be primarily a function of age. But these differences have been essentially eliminated by the fact that the weight-height values were derived by age.

In combining the values by age, irrespective of race, a further adjustment was made, necessitated by the fact that the racial composition of the examined

registrants was different from that of all registrants or of the Census population. The additional adjustment was made by weighting the adjusted age-race values of the examined registrants by the Census ratios (Table 418). These adjusted values are given in Table 411, marked "Total."

The weight-height values presented here may therefore be considered as generally relating to men within the (18–37) age-range.

It will be noted that the correlation coefficients of weight on height (Table 411) tend to decrease with age for both white and Negro men, within the given (18–37) age-range. In other words, weight becomes patently less a function of height as age increases. The correlation coefficients are higher for Negroes than for whites, for all age groups within the (18–37) age-range.

WEIGHT-HEIGHT TABLES

Three weight-height tables (Tables 412, 414, and 416) were derived from Table 411 for men within the (18–37) age-range, separately for White, Negro, and Total (White and Negro combined). The weight-values are for each inch of height, within the height-range of 59 through 78 inches, by the following age-groups: 18–19; 20–24; 25–29; 30–34; and 35–37.

Five weight values (in pounds) are given in each table for each inch of height, by age, consisting of a midvalue (italicized) and two values on either side of the midvalue. The midvalue represents the expected average weight in pounds for individuals of the particular height within the specified age group. The two values closest to the mean (∓ 1.2816 standard errors of estimate) indicate a range of weight that is expected to include 80 per cent of the individuals of the

TABLE 412

WEIGHT FOR EACH INCH OF HEIGHT: WHITE MEN, AGES 18–37*

Height (Inches)	Weight (Pounds) According to Age					Height (Inches)	Weight (Pounds) According to Age				
	18–19	20–24	25–29	30–34	35–37		18–19	20–24	25–29	30–34	35–37
	92	96	98	99	99		119	122	124	125	126
	96	100	103	104	105		125	128	130	132	133
59	*112*	*118*	*123*	*126*	*127*	69	*146*	*151*	*156*	*159*	*160*
	132	139	148	152	153		171	177	187	192	194
	138	146	155	160	162		178	186	197	202	205
	94	98	100	101	102		122	125	127	128	129
	99	103	105	107	107		128	131	133	135	136
60	*115*	*121*	*126*	*129*	*130*	70	*150*	*154*	*160*	*163*	*164*
	135	143	151	156	157		175	182	192	196	199
	141	150	159	164	166		183	191	202	207	210
	97	101	102	104	104		126	128	130	131	132
	101	105	108	109	110		131	134	137	138	139
61	*118*	*124*	*129*	*132*	*133*	71	*154*	*158*	*164*	*166*	*168*
	139	146	155	159	161		180	186	196	201	203
	145	153	163	168	170		188	195	207	212	215

TABLE 412—(*continued*)

Height (Inches)	Weight (Pounds) According to Age					Height (Inches)	Weight (Pounds) According to Age				
	18–19	20–24	25–29	30–34	35–37		18–19	20–24	25–29	30–34	35–37
	99	103	105	106	107		129	131	133	134	135
	104	108	110	112	113		135	137	140	141	142
62	*122*	*127*	*132*	*135*	*136*	72	*158*	*162*	*168*	*170*	*172*
	142	150	159	163	164		184	191	201	205	208
	149	157	167	172	174		193	200	211	217	220
	102	105	107	109	109		132	134	136	137	138
	107	110	113	115	115		138	141	143	144	146
63	*125*	*130*	*136*	*138*	*139*	73	*162*	*166*	*172*	*174*	*176*
	146	154	162	167	168		189	195	206	210	213
	153	161	171	176	178		198	205	217	222	225
	105	108	110	111	112		136	138	139	140	142
	109	113	116	117	118		142	144	147	148	149
64	*128*	*133*	*139*	*142*	*143*	74	*166*	*170*	*176*	*178*	*181*
	150	157	166	171	172		194	200	211	215	218
	157	165	175	180	182		203	210	222	227	230
	107	111	113	114	114		139	141	143	143	145
	112	116	119	120	121		146	148	150	151	153
65	*131*	*137*	*142*	*145*	*146*	75	*170*	*174*	*180*	*182*	*185*
	154	161	170	175	177		199	205	216	220	224
	161	169	179	184	186		209	215	227	232	236
	110	113	115	117	117		143	144	146	147	148
	115	119	121	123	124		149	151	154	155	157
66	*135*	*140*	*145*	*148*	*150*	76	*175*	*178*	*184*	*187*	*189*
	158	165	174	179	181		205	210	221	225	229
	165	173	183	189	191		214	220	233	238	242
	113	116	118	119	120		147	148	150	150	152
	118	122	214	126	127		153	155	158	158	160
67	*138*	*143*	*149*	*152*	*153*	77	*180*	*183*	*189*	*191*	*194*
	162	169	178	183	185		210	215	226	231	234
	169	177	188	193	195		220	226	238	243	247
	116	119	121	122	123		151	151	153	154	156
	121	125	127	129	130		157	159	161	162	164
68	*142*	*147*	*153*	*155*	*157*	78	*184*	*187*	*193*	*195*	*199*
	166	173	183	187	189		216	220	232	236	240
	174	182	192	198	200		225	231	244	249	253

* The weight values were computed from the respective regression equations of the logarithms of weight (log$_{10}$ W—in pounds) on height (H—in inches), by age. The italicized values represent the expected mean values of weight for the particular height. The values given above and below the mean values represent ∓1.2816 and ∓1.6449 standard error of estimate, covering within this range 80 and 90 per cent of the population around the mean, respectively. The first range is expected thus to eliminate 10 per cent and the other 5 per cent of the population of the particular height, on either side of the given range. The basis for these values are height and weight data on white registrants processed for military service during World War II (January 1943 through January 1944). The examined registrants were measured without their shoes and were weighed with their clothing removed.

TABLE 414

WEIGHT FOR EACH INCH OF HEIGHT: NEGRO MEN, AGES 18–37*

Height (Inches)	Weight (Pounds) According to Age					Height (Inches)	Weight (Pounds) According to Age				
	18–19	20–24	25–29	30–34	35–37		18–19	20–24	25–29	30–34	35–37
	96	99	99	99	99		123	127	128	128	128
	99	103	104	104	104		128	132	134	134	134
59	*113*	*118*	*120*	*122*	*124*	69	*146*	*151*	*155*	*157*	*159*
	128	134	139	143	147		166	172	179	185	188
	133	139	145	150	154		172	179	187	193	198
	98	102	102	102	102		127	131	132	131	131
	102	106	106	107	107		131	136	137	137	137
60	*116*	*121*	*123*	*125*	*127*	70	*149*	*155*	*159*	*161*	*163*
	132	138	143	147	150		170	176	184	189	193
	137	143	149	154	158		176	183	192	198	203
	101	104	105	104	104		130	134	135	134	134
	104	108	109	109	110		135	139	141	141	141
61	*119*	*124*	*126*	*128*	*130*	71	*153*	*159*	*163*	*165*	*167*
	135	141	146	151	154		175	181	189	194	198
	140	146	152	158	162		181	188	197	203	208
	103	107	107	107	107		133	137	138	138	138
	107	111	112	112	112		138	143	144	144	144
62	*122*	*127*	*129*	*132*	*133*	72	*157*	*163*	*167*	*169*	*171*
	139	145	150	155	158		179	186	193	199	203
	144	150	156	162	166		186	192	202	208	213
	106	110	110	110	110		137	141	142	141	141
	110	114	115	115	115		142	146	148	148	148
63	*125*	*130*	*133*	*135*	*137*	73	*161*	*167*	*171*	*174*	*176*
	142	148	154	159	162		184	190	198	204	208
	148	154	160	166	170		191	197	207	214	218
	109	113	113	113	113		140	144	146	145	145
	113	117	118	118	118		145	150	152	152	152
64	*128*	*133*	*136*	*138*	*140*	74	*166*	*171*	*176*	*178*	*180*
	146	152	158	163	166		188	195	203	209	213
	151	158	164	170	174		195	202	212	219	224
	111	115	116	116	115		144	148	149	149	148
	116	120	121	121	121		149	154	156	156	156
65	*132*	*137*	*140*	*142*	*144*	75	*170*	*175*	*180*	*183*	*185*
	150	156	162	167	170		193	200	209	215	219
	155	162	169	175	179		200	207	218	225	230
	114	118	119	118	118		148	152	153	153	152
	119	123	124	124	124		153	157	160	160	160
66	*135*	*140*	*143*	*146*	*147*	76	*174*	*180*	*185*	*188*	*189*
	154	160	166	171	175		198	205	214	220	224
	159	166	172	179	183		206	213	223	230	235

TABLE 414—(*continued*)

Height (Inches)	Weight (Pounds) According to Age					Height (Inches)	Weight (Pounds) According to Age				
	18–19	20–24	25–29	30–34	35–37		18–19	20–24	25–29	30–34	35–37
67	117	121	122	122	121	77	151	156	157	156	156
	122	126	127	127	127		157	161	164	164	164
	138	*144*	*147*	*149*	*151*		*179*	*184*	*190*	*192*	*194*
	158	164	170	175	179		203	210	220	226	230
	163	170	178	184	188		211	218	229	236	241
68	120	124	125	125	124	78	155	159	161	160	160
	125	129	130	130	131		161	166	168	168	168
	142	*147*	*151*	*153*	*155*		*183*	*189*	*195*	*197*	*199*
	162	168	175	180	184		209	215	225	232	236
	168	174	182	188	193		216	223	235	242	248

* See Footnote to Table 412.

particular height. In other words, 40 per cent of these individuals are expected to fall within this weight-range below the mean weight, and 40 per cent above the mean weight, with 10 per cent falling outside of either of these ranges. The other two weight values, farther removed from mean value (∓ 1.6449 standard errors of estimate), indicate an expected coverage of 90 per cent of the particular height-group. (See footnote to Table 412.)

These are, of course, arbitrary ranges of weight values not an attempt to define "normal weight," "underweight," or "overweight." In this respect, these ranges of values resemble those in the Medico-Actuarial Mortality investigation of 1909–1912. The latter ranges are values arbitrarily chosen as 20 per cent under or over the observed average weight [1], although it has become commonplace to consider the values within these ranges as "normal," and those outside these ranges as "underweight" (more than 20 per cent under the average weight) and "overweight" (more than 20 per cent over the average weight).

It should be kept in mind that the weight and height measurements presented here were made in the nude, while the applicants for insurance involved in the 1909–1912 study were measured in their shoes and weighed in their "ordinary indoor clothes, with coats sometimes removed" [1].

In this connection the following facts might be mentioned:

(a) The average height of the examined registrants of World War II was 68.0 inches and their average weight was 150.5 pounds.

(b) The average height of the inductees of World War II was 68.1 inches and their average weight was 150.5 pounds.

(c) The inductees of World War II were on the average 0.6 inches taller and 8.9 pounds heavier than those of World War I [3, 4].

APPENDIX: TECHNICAL NOTE

As stated in the text, adjustments for disproportionality were made in deriving the coefficients of correlation for the examined groups (inducted plus

TABLE 416

WEIGHT FOR EACH INCH OF HEIGHT, BY AGE: MEN, AGES 18–37*

Height (Inches)	Weight (Pounds) According to Age					Height (Inches)	Weight (Pounds) According to Age				
	18–19	20–24	25–29	30–34	35–37		18–19	20–24	25–29	30–34	35–37
	92	96	98	99	99		119	122	124	125	126
	96	101	103	104	105		125	128	131	132	133
59	*112*	*118*	*123*	*126*	*126*	69	*146*	*151*	*156*	*159*	*160*
	131	139	147	151	153		170	177	187	191	193
	137	146	155	159	161		178	185	196	201	204
	95	99	100	101	102		123	125	127	128	129
	99	103	105	107	107		128	131	134	135	136
60	*115*	*121*	*126*	*129*	*129*	70	*149*	*154*	*160*	*162*	*164*
	135	142	150	155	156		174	181	191	195	198
	141	149	158	163	165		182	190	201	206	209
	97	101	103	104	104		126	129	130	131	132
	101	106	108	109	110		131	135	137	138	139
61	*118*	*124*	*129*	*132*	*133*	71	*153*	*158*	*164*	*166*	*168*
	138	146	154	158	160		179	186	196	200	203
	144	153	162	167	169		187	195	206	211	214
	100	103	105	106	107		129	132	134	134	135
	104	108	111	112	112		135	138	140	141	143
62	*121*	*127*	*132*	*135*	*136*	72	*157*	*162*	*168*	*170*	*172*
	142	150	158	162	164		184	190	200	205	208
	148	157	166	171	173		192	199	211	216	219
	102	106	108	109	109		132	135	137	137	138
	107	111	113	115	115		138	141	144	145	146
63	*125*	*130*	*135*	*138*	*139*	73	*162*	*166*	*172*	*174*	*176*
	146	153	162	166	168		189	195	205	210	213
	152	160	170	175	177		197	204	216	221	224
	105	109	110	111	112		136	138	140	141	142
	110	114	116	117	118		142	145	147	148	150
64	*128*	*133*	*139*	*141*	*142*	74	*166*	*170*	*176*	*178*	*180*
	149	157	166	170	172		194	200	210	215	218
	156	164	174	179	181		202	209	221	226	230
	108	111	113	114	115		140	142	143	144	145
	113	116	119	120	121		146	148	151	152	153
65	*131*	*137*	*142*	*145*	*146*	75	*170*	*174*	*180*	*182*	*185*
	153	161	170	174	176		199	205	215	220	223
	160	168	178	183	186		207	214	226	232	235
	111	114	116	117	117		143	145	147	147	149
	115	119	122	123	124		150	152	155	155	157
66	*135*	*140*	*145*	*148*	*149*	76	*175*	*178*	*185*	*187*	*189*
	157	165	174	178	180		204	210	221	225	228
	164	172	183	188	190		213	220	232	237	241

TABLE 416—(*continued*)

Height (Inches)	Weight (Pounds) According to Age					Height (Inches)	Weight (Pounds) According to Age				
	18–19	20–24	25–29	30–34	35–37		18–19	20–24	25–29	30–34	35–37
	113	117	119	119	120		147	149	151	151	152
	118	122	125	126	127		154	156	158	159	161
67	*138*	*144*	*149*	*151*	*153*	77	*179*	*183*	*189*	*191*	*194*
	161	169	178	182	185		209	215	226	230	234
	169	177	187	192	195		219	225	238	243	247
	116	120	121	122	123		151	152	154	154	156
	122	125	128	129	130		158	159	162	163	164
68	*142*	*147*	*152*	*155*	*156*	78	*184*	*187*	*194*	*196*	*198*
	166	173	182	187	189		215	220	231	236	240
	173	181	192	197	199		224	231	243	248	253

* See footnote to Table 412.

disqualified registrants), by age and race, and by age alone (irrespective of race). These adjustments were necessitated by the following facts. First, no copies of the physical examination forms were received for Navy inductees. Consequently, direct additions of the bivariate distributions of the inductees by height and weight to the bivariate distributions of disqualified registrants would have given undue weight to the disqualified subgroups and would thus have introduced a bias in the calculations, since differences were found between the weight-height values of the subgroups. Second, the racial composition of the examined registrants was different from that of all registrants or the Census population for that period. Since there were differences in the height and weight

TABLE 417

RATIOS OF INDUCTED TO DISQUALIFIED IN SAMPLE VERSUS EXPECTED RATIOS OF QUALIFIED TO DISQUALIFIED

Age	White				Negro			
	Sample[1]		Expected[2]		Sample[1]		Expected[2]	
	Inducted	Disqualified	Qualified	Disqualified	Inducted	Disqualified	Qualified	Disqualified
18–19	.7052	.2948	.7593	.2407	.5683	.4317	.5931	.4069
20–24	.6935	.3065	.7491	.2509	.5470	.4530	.5721	.4279
25–29	.5856	.4144	.6508	.3492	.4135	.5865	.4385	.5615
30–34	.5025	.4975	.5712	.4288	.3316	.6684	.3546	.6454
35–37	.4160	.5840	.4844	.5156	.2830	.7170	.3042	.6958

[1] Based on Table 409.

[2] Derived from reports ("Monthly Report of Examination of Joint Army and Navy Inductees") submitted monthly during World War II to the Surgeon General's Office by each induction station. These reports showed, by race, the number of registrants examined, number disqualified (by disqualifying cause), and the number of inductees by type of qualification and department of assignment: Army or Navy.

values by race, additional adjustments were required for this disproportionality in order to obtain values as close as possible representative of the total population (white plus Negro).

The adjustments for disproportionality were computed in the following manner:

(a) *Examined group (inducted plus disqualified), by age and race.* The data were adjusted by weighting the computed separate values (means, standard deviations, and covariances) of these subgroups by the "expected" ratios as given in Table 417, by age and race.

(b) *Examined group (white plus Negro), by age.* The data were adjusted by further weighting the adjusted values derived in (a) by the white to Negro ratios given in Table 418, by age.

The computations of the weighted values (means, standard deviations, and covariances) were as follows:

Let $\overline{X}_i$ and $\overline{Y}_i$ represent the respective height and weight means of a subgroup, and $\overline{X}$ and $\overline{Y}$ the mean values for the corresponding total group. Obviously,

$$\overline{X} = \sum w_i \overline{X}_i \qquad (i = 1, \cdots, k) \tag{1}$$

where w_i is the proportion of the subgroup within its total group ($\sum w_i = 1.0$). For the (20–24) white age group, for instance, the w_i's are .7491 and .2509 (Table 417), when combining the respective inducted and disqualified subgroups. For the same age group, the w_i's are .8996 and .1004 (Table 418), when combining the corresponding white and Negro subgroups. The same relationship (1) holds, of course, for $\overline{Y}$.

The variance $V(X)$ of the total group is:

$$V(X) = \sum w_i[V(X_i) + \overline{X}_i^2] - \overline{X}^2 \qquad (i = 1, \cdots, k) \tag{2}$$

where $V(X_i)$ is the variance of the subgroup; $\overline{X}$, $\overline{X}_i$ and w_i are the same as in (1), above. The same holds for $V(Y)$. (See: Kenney [6], p. 101, corollary 1.)

TABLE 418

RATIOS OF WHITE TO NEGRO IN THE ADJUSTED SAMPLE VERSUS THE RATIOS OF WHITE TO NEGRO IN THE MALE POPULATION, BY AGE

Age	Ratios			
	Adjusted Sample		Male Population*	
	White	Negro	White	Negro
18–19	.8875	.1125	.8950	.1050
20–24	.8422	.1578	.8996	.1004
25–29	.8314	.1686	.9047	.0953
30–34	.8254	.1746	.9049	.0951
35–37	.8279	.1721	.9096	.0904

* Based on special communication from the Bureau of the Census relating to 1943.

Following essentially the procedures used by Kenney [6] in deriving weighted variance (2), the following relationship was established for weighted covariance (Cov):

$$\text{Cov}\ (XY) = \sum w_i[\text{Cov}\ (X_iY_i) + \overline{X}_i\overline{Y}_i] - \overline{XY} \quad (i = 1, \cdots, k) \quad (3)$$

where Cov (X_iY_i) is covariance of the subgroup. The other values have been explained in (1), above.

ACKNOWLEDGMENTS

The author is indebted to the many persons in the Medical Statistics Division, Office of The Surgeon General, Army, who were involved in this study, and, especially, to Betty Feldman, Lillian Smith, Henrietta Gingrich, Josephine Balthazar, Eloise LaRue, and their coworkers, who carried the brunt of the numerous required computations. Special acknowledgment is made to Harold A. Grossman for accomplishing the more complex computations, and, specifically, for deriving "weighted covariance" (Technical Note).

REFERENCES

[1] Association of Life Insurance Medical Directors and Actuarial Society of America, *Medico-Actuarial Mortality Investigation*, 1. 1912.

[2] Brozek, Joseph, and Keys, Ancel, "Limitations of the normal body weights as a criterion of normality," *Science* 112 (1950), 788.

[3] Davenport, Charles B., and Love, Albert J., *Army Anthropology, Vol. 15, Part One.* Washington: Medical Department, U. S. Army, World War, 1921.

[4] Karpinos, Bernard D., "Height and Weight." (Chapter in *General Characteristics and Fitness of Selective Service Registrants for Military Service, World War II*, a projected volume in History of the Medical Department, U. S. Army, World War II.)

[5] Kendall, Maurice G., *The Advanced Theory of Statistics, Vol. II.* London: Charles Griffin and Company Limited, 1946.

[6] Kenney, John F., *Mathematics of Statistics, Part One.* New York: D. Van Nostrand Co., Inc., 1951.

[7] Kenney, John F., and Keeping, E. S., *Mathematics of Statistics, Part Two.* New York. D. Van Nostrand Co., Inc., 1953.

[8] Newman, Russell W., *The Assessment of Military Personnel by 1912 Height-Weight Standards.* Quartermaster Climatic Research Laboratory, Army, Report No. 194, 1952.

[9] O'Brien, Ruth, Girshick, Meyer A., and Hunt, Eleanor P., *Body Measurements of American Boys and Girls for Garment and Pattern Construction.* Washington: U. S. Department of Agriculture, Miscellaneous Publication No. 366, 1941.

[10] Rosenbaum, S., "Heights and weights of the Army intake, 1951," *J. Royal Stat. Soc.*, 117 (1954), 331–347.

[11] War Department, *Standards of Physical Examination During Mobilization, MR(1–9)*, 1942.

ESTIMATION OF SURVIVORSHIP IN CHRONIC DISEASE: THE "ACTUARIAL" METHOD*

LILA ELVEBACK
Tulane University

In most follow-up studies in chronic disease assumptions on the parametric form of the survivorship function are not justified. A simple frequency estimate, arising from polygonal approximation, is discussed. It is shown that, for proper limitation of the period over which the survivorship is to be estimated, bias due to nonlinearity may be controlled, without loss in efficiency, by proper choice of the sub-intervals. It is concluded that the method proposed is appropriate and highly efficient for large scale follow-up studies such as those carried out by the Mayo Clinic.

THE problem of estimation of the survivorship function in follow-up studies in chronic disease is characterized by the fact that some part of the sample observations are not complete over the entire period under study, and also by the necessity of avoiding an assumption concerning the parametric form of the survivorship function.

The accepted and almost universally used method of dealing with this situation has consisted in the subdivision of the total period into a set of small sub-intervals, and the estimation of the conditional probability of survival over each of the sub-intervals. This method, which has been in use for many years, has been called the "life table" or "modified life table" method. In 1950, it was designated as "actuarial" by Berkson [3], in order to distinguish it from a direct calculation frequently employed by clinicians. Since that time this designation has gained fairly wide acceptance in the literature on follow-up studies [2, 4, 13, 14, 16, 17], and will be adopted here.

Recently a variety of alternative methods of estimating the survivorship function have been proposed [13, 14, 17]. Some of these are based on assumptions concerning the functional form of

$$P(t) = \text{Prob (survival to time } t \text{ following entry)}$$

over the entire period under study, and are unsuitable for use in most chronic disease studies. Others, while completely non-parametric in this respect, involve clerical and computational work which becomes prohibitive in large studies.

It has been pointed out that actuarial methods are inefficient alternatives to the one period estimator when the form of the survivorship function is assumed to be known. This is true, but does not detract from the appropriateness of the method for the majority of chronic disease studies in which the form of the survivorship function is, in fact, unknown.

It is the purpose of this paper to propose a simple actuarial estimate based solely on the interval frequencies and to illustrate its performance. The conclusion is reached that the method proposed constitutes a very satisfactory

* Most of this work was carried out during a Fellowship in the Section of Biometry and Medical Statistics of the Mayo Clinic. Acknowledgement is due to Dr. Joseph Berkson for generous help in all phases of the work.

procedure for estimation of the survivorship function over a large class of chronic disease studies.

For the purposes of this paper the actuarial method is taken to be one which is appropriate when the functional form of $P(t)$ is unknown, and in which

1. the period $(0, T)$ under study is partitioned into a number of sub-intervals

$$(0, t_1), \cdots, (t_{a-1}, t_a),$$

where $t_a = T$.

2. the parameters to be estimated are the conditional probabilities

$$1 - q_k = \text{Prob (survival to } k+1 \mid \text{survival to } k)$$

for $k = 0, 1, \cdots, a-1$.

3. $P(t_k)$, for $k = 1, \cdots, a-1$, is estimated as

$$\prod_{i=0}^{k-1} (1 - \bar{q}_i),$$

where $\bar{q}_i$ is the estimate of q_i.

The method presented involves, in addition to the properties above, the approximation of the unknown $P(t)$ by the polygonal function

$$P(t) = 1 - q_0 t \quad \text{for} \quad 0 < t \leq 1,$$

and by

$$P(t) = [1 - (t-k)q_k] \prod_{i=0}^{k-1} (1 - q_i) \text{ for } k \leq t \leq k+1 \text{ and } k = 1, 2, \cdots, a-1.$$

Linearity over the sub-intervals is chosen as a natural first approximation to the unknown $P(t)$. In general the lengths of the intervals should be chosen in such a way as to insure the adequacy of the linear approximation, and in most studies intervals of unequal lengths are suggested. It will be seen from the results below that the common practice of using a calendar year for each interval will lead to satisfactory results in many studies. For notational convenience and simplicity of the numerical examples, the intervals here have been chosen to be of equal length. In general, the order of the approximation may be controlled by choice of the number and lengths of the intervals.

The estimates suggested were developed for use in the studies conducted by the Mayo Clinic and are appropriate to the method of follow-up routinely used there. Since some of the other estimates which have been suggested [13] cannot be used for the sample information resulting when that method is used, a description of this follow-up method and an alternative method of using the sample information is necessary.

THE FOLLOW-UP PROCEDURE

The group of patient histories studied is made up of all cases in the disease-treatment group defined [7] and entering that group (at onset, diagnosis, treatment or discharge) at the clinic over some specified period. For example,

the group might consist of all patients undergoing and surviving gastric resection for the removal of malignant neoplasm of the stomach who were discharged from the hospital during the period between January 1, 1935 and December 31, 1944. For each patient the date of entry into the study is taken to be the date of his discharge from the hospital.

In some cases a single study of the group of patients is made. In the example above, if the study were initiated in January 1947, information concerning the experience of the patients subsequent to discharge would, in general, be at hand only for those patients who had returned to the clinic for treatment in the interval between discharge and the initiation of the study. The follow-up procedure would consist of sending letters of inquiry to each member of the group, with the exception of course, of those who were known to have died before the opening of the inquiry. The process of mailing the letters might cover a period of several weeks. The period during which replies would be received, second letters written, and further attempts to trace the patient undertaken might cover from three months to a year. At the conclusion of this period the members of the original group would fall into the following classes:[1]

a_1) those reported to be dead, date of death known

b_1) those who reply and are reported living at some date during the inquiry period.

In some cases the follow-up is brought up to date every few years. Further inquiry might be made in 1953 concerning the subsequent experience of the *group* b_1 above. The 1953 study might include, in addition the new cases entering the group in the interval from January 1, 1945 through December 31, 1950. In the 1953 inquiry, letters would be sent to members of the group b_1 and to those entering in 1945–1950. At the close of the new inquiry the total group is classified as:

a_2) those reported to be dead, date of death known

b_2) those reported alive at some date during the 1953 inquiry.

Two methods of using the sample information are described below.

The Common Closing Date[2]

The follow-up method described above results in observational periods which differ from individual to individual not only in length but also in the calendar date of last report. The dates of last report, with very few exceptions, fall within the inquiry period and may be spread out over most of a calendar year. The common closing date method[3] of using the sample information corresponds to choice of a single calendar date (which might be on or before the opening

[1] In addition there may be a few cases in which the patient is untraced or is known to have survived a number of years following entry, but for whom all attempts to discover his status during the inquiry period fail. This group is kept to a minimum in the Mayo Clinic studies, in some of which complete follow-up has been achieved for 99% of the study group. In the discussion which follows, it is assumed that untraced or partially traced histories occur in negligible instances only. If the untraced or partially traced group is large and non-randomly "selected" the results will be biased regardless of the method of estimation used.

[2] This method is not ordinarily used at the Mayo Clinic.

[3] It is to be noted that the procedure used here is not the usual application of a common closing date to information obtained from a study based on the anniversary method, in which patients are followed to an anniversary of the date of entry, but not longer than some specified date. No common closing date is used, but each patient's history is terminated at his own last anniversary of entry prior to the study date. It will be noted that if the "anniversary method" were to be used in connection with data arising from the inquiry method of follow-up described here, that the resulting loss in discarded sample information might be greater than that associated with use of the study date as a common closing date.

date of the inquiry) which is used as the closing date for the experience of every individual in the study. For each individual alive on the common closing date the information at hand concerning subsequent experience (from the common closing date to the date of last report for that individual) is discarded.

In the example described above the sample information resulting from the 1953 inquiry might be used to establish, for each patient, his status as of January 1, 1953. The longest history might cover the period between entry on January 1, 1935 and the closing date January 1, 1953 and the shortest the period between December 31, 1950 and January 1, 1953. The sample information becomes

a) the calendar date of death for those who die before January 1, 1953, which gives the *time of death* as measured from entry.
b) the *survival time* or the interval between entry date and the closing date for those who survive to the closing date, January 1, 1953.

Under this method a patient who is reported as living on August 1, 1953 and another whose date of death is known to be July 1, 1953, are both recorded as living on January 1, 1953. Some part of the observed experience of each patient who survives beyond January 1, 1953 is discarded. In the study described, since there are no entries after December 31, 1950, no information concerning survival over the first two years following entry is discarded.

In an effort to obtain a rough estimate of the relative cost in discarded information which is associated with the application of a common closing date to observations arising from the inquiry type of follow-up described above, a study was made of 1559 Mayo Clinic patients who entered the follow-up experience at the date of diagnosis of angina pectoris during the years 1940 through 1944. The common closing date was taken as February 15, 1950, the date of first reply received during the 1950 inquiry. It follows that there was no information discarded due to termination of the study during the first five years of experience. A total of 619 deaths occurred during this time. At the beginning of the sixth year following diagnosis, 940 of the patients were still alive and under observation. Of these 221 were observed to die before the common closing date, and the remaining 719 were alive on the closing date, February 15, 1950, and survived to report at some point beyond this date. A total of 264 years of observation, or an average of .36 year per person alive on February 15, 1950, was discarded by use of the common closing date. On the basis of a rough calculation the number of years of observation used in the estimation of $(q_5 \cdots, q_9)$ was 2670, and the information discarded was roughly nine per cent of that available.

In the example above in which a 1947 study of the 1935–45 entries was extended and brought up to date in 1953, it could have happened that in 1953 no information was obtainable on a more than negligible number of the individuals reported alive at the 1947 closing date. The presence of such unknowns in the follow-up would clearly preclude the use of the common closing date method as presented here.

The Date of Last Report

The inquiry method of follow-up which has been described is that ordinarily used at the Mayo Clinic. However, the common closing date method of han-

dling the sample information is not used in the analysis of the results in these inquiries. Instead all information on each member of the group available up to the last report on that individual has been included in the sample information. As stated previously, it is assumed that for patients living at last report, the date of last report is, in practically all cases, within the period of inquiry. When the inquiry is completed this information consists of two calendar dates for each person:

1) the calendar date of entry, and
2) the calendar date of death

or

1) the calendar date of entry, and
2) the latest date on which the patient is known to have been alive.

The sample information is taken to be

a) the time of death as measured from entry,

or

b) the time of withdrawal alive, which is the length of the interval between entry and the last date on which it is known that the patient was living.

The date of last report method has the considerable advantage that all available information is used. It has been suggested that the inclusion of experience subsequent to the beginning of the inquiry period will produce bias since the date of last report is, in a sense, selected by the person concerned. If so, the magnitude of the effect might be investigated empirically for each study. Since it is to be expected that the amount of bias introduced will depend on the disease in question, on the form of the inquiry letter and on many other factors, no single investigation would provide a basis for generalization. It is the opinion of the writer that the "selection bias" introduced by the inclusion of the post closing date information is a minor matter.

The Sample Information Under the Common Closing Date Method

For each person in the follow-up group let the length of history be $H_i = C - E_i$ where C is the common closing date, and E_i the entry date of the ith patient. If $H_i \leq T$, let $T_i = H_i - [H_i]$, where $[H_i]$ is the greatest integer contained in H_i. If the ith patient survives to $[H_i]$ he is referred to as a *withdrawal* in the interval $([H_i], [H_i]+1)$, and if further he survives to $H_i = [H_i] + T_i$ he is referred to as a *withdrawal alive* at time T_i in this interval.

For each person in the group let τ_i be the interval between entry and death. If $\tau_i > H_i$, its value is not observable. If $\tau_i \leq H_i$ let $t_i = \tau_i - [\tau_i]$ and the patient is referred to as a death at time t_i in the interval $([\tau_i], [\tau_i]+1)$.

Given M individuals in the total group studied. With reference to the typical interval, $(k, k+1)$, for $k = 0, 1, \cdots, a-1$, the individuals can be classified as

N_k = the number, among those for whom $H_i > k$, who survive to k.

The N_k survivors to k can be further classified as

W_k = the number, among the N_k for whom $k < H_i < k+1$
= the number of withdrawals in $(k, k+1)$
= $D_{2k} + L_k$, where
D_{2k} = the number who die before withdrawal
L_k = the number who survive to withdraw alive
$N_k - W_k$ = the number, among the N_k, for whom $H_i \geq k+1$
= the number of non-withdrawals
= $D_{1k} + S_k$ where
D_{1k} = the number who die during the interval $(k, k+1)$ among the non-withdrawals, and
S_k = the number who survive to $k+1 = N_{k+1}$.

Then $N_k = (D_{1k} + S_k) + (D_{2k} + L_k)$, and for $D_k = D_{1k} + D_{2k}$, $N_k = D_k + L_k + S_k$.

ESTIMATION OF THE INTERVAL PARAMETERS q_k

Estimates for the Common Closing Date Method

For the purposes of discussion of the relative merits of different estimates of the q_k, the quantities N_k, W_k and $T_1, \cdots, T_{W_k}$ are analogous to "sample size" and will be taken in this section as given constants. The sample observations are D_{1k} and D_{2k} and the times of death, L_k and the times of withdrawal, and $S_k = N_{k+1} = N_k - D_k - L_k$.

Since, for t on the sub-interval $k \leq t \leq k+1$,

$$\text{Prob (survival to } t \mid \text{survival to } k) = 1 - (t - k)q_k,$$

then, for any t in the interval

$$\text{Prob (death in } t, t + dt \mid \text{survival to } k) = q_k dt.$$

It follows that the contribution to the likelihood function for a death during the interval, among the $N - W$ persons whose histories are complete over the interval, is simply q_k regardless of the time of death as measured within the interval. The assumption that the survivorship function may be satisfactorily approximated by a straight line over the interval is equivalent to the assumption that the information is essentially complete when the time of death is observed only in terms of the interval in which it occurs.

The sample information for the interval is taken to consist of $N_k - W_k$ observations on a binomial variable of parameter q_k, resulting in a contribution to the likelihood function of

$$q_k^{D_{1k}}(1 - q_k)^{N_k - W_k - D_{1k}},$$

and a single observation on each of W_k binomial variables of parameters $q_k T_1, \cdots, q_k T_{W_k}$, the typical, or jth, of which results in a contribution to the likelihood function of

$$q_k T_j \text{ if death occurs between } k \text{ and } k + T_j,$$

or of

$$1 - q_k T_j \text{ in the event of survival to } k + T_j.$$

Choice of the conditional probabilities q_k as the parameters to be estimated, leads to a likelihood function which can be written in the form

$$\phi = \prod_{k=0}^{a-1} \phi_k$$

where

$$\phi_k = q_k^{D_{1k}}(1 - q_k)^{N_k - D_k - L_k}\left[\prod_{n=1}^{D_{2k}} q_k T_n\right]\left[\prod_{m=1}^{L_k}(1 - q_k T_m)\right].$$

It follows that

$$\frac{\partial^2 \log \phi}{\partial q_i \partial q_j} = 0$$

for all $i \neq j$, and the asymptotic covariance matrix of the maximum likelihood estimates $(\hat{q}_0, \cdots, \hat{q}_{a-1})$ is diagonal, so that all asymptotic covariances between the $\hat{q}_i$ are zero.

The asymptotic variance of $\hat{q}_i$ is given by

$$\left[-E\left(\frac{\partial^2 \log \phi}{\partial q_i^2}\right)\right]^{-1}.$$

For notational convenience in the discussion of the typical interval, $(k, k+1)$ and the estimation of q_k, the subscript k will be dropped. The maximum likelihood estimate,[4] $\hat{q}$, is the solution of

$$\frac{\partial \log \phi}{\partial q} = \frac{D}{q} - \frac{N - D - L}{1 - q} - \sum^{L}\left(\frac{T_m}{1 - qT_m}\right) = 0 \tag{1}$$

and must be evaluated by an iterative procedure. Further

$$-\frac{\partial^2 \log \phi}{\partial q^2} = \frac{D}{q^2} + \frac{N - D - L}{(1 - q)^2} + \sum^{L}\left(\frac{T_m}{1 - qT_m}\right)^2 \tag{2}$$

and

$$-E\left[\frac{\partial^2 \log \phi}{\partial q^2}\right] = \frac{(N - W)q + \sum^{W} qT_j}{q^2} + \frac{(N - W)(1 - q)}{(1 - q)^2} + \sum^{W}\left[\left(\frac{T_j}{1 - qT_j}\right)^2(1 - qT_j)\right]$$

$$= \frac{N - W}{q(1 - q)} + \sum^{W}\left(\frac{T_j^2}{1 - qT_j} + \frac{T_j}{q}\right),$$

and

$$[\operatorname{Var} \hat{q}]^{-1} = \frac{N - W}{q(1 - q)} + \frac{1}{q}\sum^{W}\left(\frac{T_j}{1 - qT_j}\right). \tag{3}$$

[4] This estimate was first proposed by Berkson, on the basis of a heuristic argument [4]. See Appendix, Note 2.

The clerical and computational work is greatly simplified if the estimate is based solely on the interval frequencies. Several such estimates have been suggested and used in application [1, 2, 3, 4, 8, 9, 13, 16, 18]. A discussion of their properties must be carried out within the framework of a model which includes, in addition to an assumption on, or an approximation to, $P(t)$ over the interval, information concerning the distribution of the T_i for the withdrawals. No direct set of observations on the T_i under the common closing date method has been made in connection with the Mayo Clinic data. However, for the date of last report method the T_i for the withdrawals alive have been determined for some 1250 cases, of which 450 were taken from a study of cancer of the stomach and the remainder from a study of angina pectoris. For each disease several typical intervals were used. The resulting set of six distributions of times of withdrawal alive showed no marked or systematic departure from that expected under a uniform distribution of the T_i values for all withdrawals, which is given by

$$f(T) = \frac{1}{1 - q/2} - \frac{q}{1 - q/2} T \quad \text{on } 0 \leq T \leq 1.$$

Use, as a averaging device, of the assumption that the T_i arise as observations on a uniformly distributed random variable will be made in the following discussion. It is believed that this assumption will serve as an adequate approximation in most follow-up studies of the type under discussion.

This simplifying assumption gives, for the W withdrawals

$$\text{Prob (death before withdrawal time)} = q/2.$$

If the T_i observations are censored, and this averaging device is adopted, then the sample information is regarded as

$N - W$ observations on a binomial variable of parameter q,

and

W observations on a binomial variable of parameter $q/2$.

Then

$$\phi = q^{D_1}(1 - q)^{N-D-L}\left(\frac{q}{2}\right)^{D_2}\left(1 - \frac{q}{2}\right)^{L}$$

and the maximum likelihood estimate corresponding to this simplified set of observations is[5]

$$\hat{q}_L = \frac{1}{2N}\left[2N - L + D - \sqrt{(2N - L + D)^2 - 8DN}\right]. \tag{4}$$

Again the asymptotic variance is given by

$$\left[-E\left(\frac{\partial^2 \log \phi}{\partial q^2}\right)\right]^{-1}.$$

[5] This estimate was also proposed by Berkson as an approximation to the solution of (1).

Here

$$-\frac{\partial^2 \log \phi}{\partial q^2} = \frac{D}{q^2} + \frac{N - D - L}{(1-q)^2} + \frac{L}{(2-q)^2},$$

$$-E\left[\frac{\partial^2 \log \phi}{\partial q^2}\right] = \frac{(N-W)q + W\frac{q}{2}}{q^2} + \frac{(N-W)(1-q)}{(1-q)^2} + \frac{\frac{W}{2}(2-q)}{(2-q)^2}$$

and

$$\text{Var } \hat{q}_L = \frac{q(1-q)(2-q)}{N\left(2 - q - \frac{W}{N}\right)}.$$

It is of interest to include in the comparison the frequency estimate

$$\hat{q}_W = \frac{D}{N - \frac{W}{2}}$$

discussed by Littell [13]. Under the same assumptions used above, since D is the sum of two independent binomial variables,

$$\text{Var } \hat{q}_W = \frac{\text{Var } D_1 + \text{Var } D_2}{(N - W/2)^2} = \frac{(N-W)q(1-q) + W\frac{q}{2}\left(1 - \frac{q}{2}\right)}{(N - W/2)^2}.$$

To evaluate the loss in efficiency due to discarding the T_i values for the withdrawals alive, the variance of the maximum likelihood estimate corresponding to the use of the entire set of observable random variables has been evaluated by use of the same simplifying assumption on the distribution of the T_i. From (2) and (3)

$$[\text{Var } \hat{q}]^{-1} = \frac{N-W}{q(1-q)} + \frac{W}{q}\int_0^1 \frac{T}{1-qT}\,dT$$

$$= \frac{N-W}{q(1-q)} - \frac{W}{q^2}[q + \log(1-q)].$$

Under linear survivorship the asymptotic efficiency of the frequency estimates is high, even for large q and a large proportion of withdrawals. For example, if $q = .2$ and $W = .2N$, the efficiency of $\hat{q}_L$ and $\hat{q}_W$ are .987 and .984 respectively. If $q = .5$ and $W = .5N$, the corresponding efficiencies are .962 and .927. Note that the use of $\hat{q}_W$ is possible only if the common closing date is used.

Estimation Under the Date of Last Report Method

If all of the available information is used, the H_i values are no longer defined. For those who are alive at the date of last report, H_i' may be taken as the in-

terval between entry and date of withdrawal alive. If $H_i' \leq T = a$, then for some k

$$H_i' = k + T_i \quad \text{where} \quad 0 < T_i \leq 1$$

and the individual is referred to as a withdrawal alive at T_i in the $(k+1)$st interval, and is a member of the group L_k. For those whose date of death is known, no similar "observation time" or "length of history" is defined. The sample information for the typical interval consists of N, D, L and $T_1, \cdots, T_L$. The frequency W is no longer observable and the D_1 and D_2 groups are no longer distinguishable. On the other hand, the information required for the use of either $\hat{q}$ of (1) or $\hat{q}_L$ of (4) is available. However the evaluation of the asymptotic variances cannot be carried out without further assumptions on the lengths of history or observation times. In the model discussed below the common closing date method and the date of last report method correspond to different systems of censoring of the sample information.

A GENERAL MODEL

For each member of the total group, the length of history, H, is taken to be an observation on a random variable, independent of the time of death, and with distribution given by

$$P(H > t) = 1 - \rho_0 t \quad \text{for} \quad 0 < t \leq 1,$$

and by

$$P(H > t) = [1 - (t - k)\rho_k] \prod_{i=0}^{k-1} (1 - \rho_i) \text{ for } k < t \leq k + 1 \text{ and } k = 1, \cdots, a - 1.$$

Then

$$\rho_k = \text{Prob } (H < k + 1 \mid H > k).$$

The ρ_k are parameters to be estimated. The time of death τ, is governed by the survivorship function

$$P(t) = 1 - q_0 t \quad \text{for} \quad 0 < t \leq 1$$

and

$$P(t) = [1 - (t - k)q_k] \prod_{i=0}^{k-1} (1 - q_i) \text{ for } k < t \leq k + 1 \text{ and } k = 1, \cdots, a - 1.$$

and is observable only in those cases in which $\tau < H$.

The probability,

$$P(t) = \text{Prob (survival to } t\text{), must be distinguished from}$$

$$Q(t) = \text{Prob (survival to } t \text{ and under observation at } t\text{)}$$

$$= \text{Prob } (H > t \text{ and survival to } t).$$

Fix and Neyman [10], in describing a competitive risk model[9] for use in chronic disease studies, refer to probabilities such as $P(t)$ as "net" and to probabilities such as $Q(t)$ as "crude." Note that

$$P(k) = \prod^{k-1} (1 - q_i),$$

whereas the expected number of persons alive and under observation at k, or

$$E(N_k) = M \prod^{k-1} (1 - q_i)(1 - \rho_i) = MQ(k).$$

Under the general model, given survival under observation to k the probabilities of the observable events are, for $k \leq t \leq k+1$

$$\text{Prob (withdrawal alive in } t, t + dt) = \rho_k(1 - q_k t)dt$$

$$\text{Prob (death observed in } t, t + dt) = q_k(1 - \rho_k t)dt.$$

Again the choice of the conditional probabilities q_k and ρ_k as the parameters to be estimated, leads to a likelihood function which can be written in the form

$$\phi = \prod_{k=0}^{n-1} \phi_k,$$

where

$$\phi_k = [(1 - \rho_k)(1 - q_k)]^{S_k} \left[\prod_{m=1}^{D_k} q_k(1 - \rho_k t_m) \right] \left[\prod_{m=1}^{L_k} \rho_k(1 - q_k T_m) \right].$$

ϕ_k can also be factored into functions of ρ_k and q_k separately. It follows that all asymptotic covariances between estimates $(\hat{q}_0, \cdots, \hat{q}_{n-1}, \hat{\rho}_0, \cdots, \hat{\rho}_{n-1})$ are zero, and again the reciprocal of the asymptotic variance of $\hat{q}_k$ is

$$-E\left(\frac{\partial^2 \log \phi}{\partial q_k^2}\right).$$

For the typical interval, dropping the subscript k, and for N persons alive and under observation at the beginning of the interval, the maximum likelihood estimate $\hat{q}$ is the solution of

$$\frac{\partial \log \phi}{\partial q} = \frac{D}{q} - \frac{N - D - L}{1 - q} - \sum_{m=1}^{L} \left(\frac{T_m}{1 - qT_m} \right) = 0$$

[9] The model described above is equivalent to a competitive risk model in which the intensities of risk of death and of withdrawal alive, $r_D(t)$ and $r_L(t)$ respectively, are, for $k \leq t \leq k+1$, given by

$$r_D(t)dt = \text{Prob (death in } t, t + dt | \text{survival to } t) = \frac{q_k}{1 - q_k t} dt,$$

and

$$r_L(t)dt = \text{Prob } (t < H < t + dt | H > t) = \frac{\rho_k}{1 - \rho_k t} dt.$$

The quantity $r_D(t)$ is also referred to as the "force of mortality" and is related to the net survivorship by

$$r_D(t) = -\frac{d}{dt} \log P(t).$$

which corresponds exactly to the solution of (1). Also

$$-E\left(\frac{\partial^2 \log \phi}{\partial q^2}\right) = E\left\{\frac{D}{q^2} + \frac{N-D-L}{(1-q)^2} + \sum_{m=1}^{L}\left(\frac{T_m}{1-qT_m}\right)^2\right\}$$

$$= N\left\{\frac{1-\rho/2}{q} + \frac{1-\rho}{1-q} + \rho\left(1-\frac{q}{2}\right)\int_0^1 \left(\frac{T}{1-qT}\right)^2 \frac{\rho(1-qT)}{\rho(1-q/2)}\,dT\right\},$$

and

$$(\text{Var } \hat{q})^{-1} = N\left\{\frac{1-\rho}{q(1-q)} - \frac{\rho}{q^2}\left[q + \log(1-q)\right]\right\}.$$

Finally, removing the condition on N_k persons alive and under observation at time k, and replacing N_k by its expectation.

$$(M \text{ Var } \hat{q}_k)^{-1} = \left\{\frac{1-\rho_k}{q_k(1-q_k)} - \frac{\rho_k}{q_k^2}\left[q_k + \log(1-q_k)\right]\right\} \prod_{i=0}^{k-1}(1-q_i)(1-\rho_i).$$

For the typical interval, the maximum likelihood estimate of ρ is the solution of

$$\frac{\partial \log \phi}{\partial \rho} = \frac{L}{\rho} - \frac{N-D-L}{1-\rho} - \sum^{D}\left(\frac{t_n}{1-\rho t_n}\right) = 0.$$

Frequency Estimates

For the frequency estimates of q and ρ of the typical interval, since

$$\text{Prob (death observed during the interval)} = q(1-\rho/2)$$

$$\text{Prob (withdrawal alive in the interval)} = \rho(1-q/2),$$

the joint sample probability becomes

$$\phi = [(1-\rho)(1-q)]^S[q(1-\rho/2)]^D[\rho(1-q/2)]^L.$$

The solution of $d \log \phi/dq = 0$ gives, as the maximum likelihood estimate of q, the quadratic estimate $\hat{q}_L$ of (4). Similarly

$$\hat{\rho}_L = \frac{1}{2N}\left[2N - D + L - \sqrt{(2N-D+L)^2 - 8LN}\right]. \tag{5}$$

Also

$$-E\left(\frac{\partial^2 \log \phi}{\partial q^2}\right) = E\left\{\frac{D}{q^2} + \frac{N-D-L}{(1-q)^2} + \frac{L}{(2-q)^2}\right\}$$

$$= N\left\{\frac{1-\rho/2}{q} + \frac{1-\rho}{1-q} + \frac{\rho/2}{2-q}\right\}$$

and, for N given

$$\text{Var } \hat{q}_L = \frac{q(1-q)(2-q)}{N(2-q-\rho)}.$$

Again, removing the condition that N_k is given, and denoting by $\hat{q}_{L_k}$ the quadratic estimate of q_k

$$\text{Var } \hat{q}_{L_k} = \frac{q_k(1-q_k)(2-q_k)}{(2-q_k-\rho_k)E(N_k)} \quad \text{where} \quad E(N_k) = M \prod_{i=0}^{k-1}(1-q_i)(1-\rho_i). \quad (6)$$

The Effect of Non-Linearity of the Survivorship Function

In general application to large studies in chronic disease, in which the form of the survivorship function is unknown, the actuarial method is taken to be the method of choice. The question as to how the $(q_0, \cdots, q_{a-1})$ shall be estimated remains. With reference to the choice of the frequency estimate $\hat{q}_L$ of the linear model the following questions arise:

1. How sensitive is $\hat{q}_L$ to non-linearity of $P(t)$ over the sub-intervals?
2. What is the cumulative effect of the bias of $\hat{q}_L$ on the estimate $\hat{P}(T)$ when $P(t)$ is non-linear?
3. What is the effect on the variance of $\hat{P}(T)$ of using a larger number of smaller intervals in order to decrease the bias?

Concerning the last of these questions, Littell [13] discusses an actuarial estimate of $P(T)$, namely one based on $\hat{q}_W$, for a model in which the functional form of $P(t)$ is known to be exponential, and for which $P(t)$ is estimated over a period $(0, T)$ on which all of the available histories become incomplete. He concludes that the efficiency of the actuarial estimate is low, and that subdivision into many small subintervals is to be avoided on the grounds that the desired decrease in the bias of $\hat{P}(T)$ is accompanied by a rapid increase in its variance.

Since in many chronic disease studies the assumption of an exponential $P(t)$ over the entire period $(0, T)$ may be inadequate, it is of more interest to compare the performance of the proposed actuarial estimate based on $\hat{q}_L$ with the optimal estimate for the case in which $P(t)$ consists of a series of exponential arcs over the sub-intervals. To avoid extrapolation the period $(0, T)$ is restricted so that some histories complete over the entire period are available.

The Alternative Model

Following the method of Fix and Neyman [10] and taking here for the alternative model,

$$r_D(t) = \mu \quad \text{and} \quad r_L(t) = \frac{\rho}{1-\rho t}$$

the required probabilities become, for the typical interval, and $0 \leq t \leq 1$,

$$\text{Prob (survival under observation to } t) = (1-\rho t)e^{-\mu t}$$

$$\text{Prob (death before } t) = (1-e^{-\mu t})(1-\rho/\mu) + \rho t e^{-\mu t}$$

$$\text{Prob (withdrawal alive before } t) = \frac{\rho}{\mu}(1-e^{-\mu t})$$

$$\text{Prob (death observed in } t, t+dt) = (1-\rho t)\mu e^{-\mu t}dt$$

$$\text{Prob (withdrawal alive in } t, t+dt) = \rho e^{-\mu t}dt.$$

The maximum likelihood estimate of q for the typical interval is

$$\hat{q} = 1 - e^{-\hat{\mu}} \quad \text{where} \quad \hat{\mu} = \frac{D}{S + \sum^{L} T_m + \sum^{D} t_n} \tag{7}$$

and the asymptotic variance is

$$\text{Var } \hat{q} = \frac{(1-q)^2\mu^2}{E(D)}$$

where $E(D)$ is the expected value of D, and for N given

$$E(D) = N\left[\rho(1-q) + q\left(1 - \frac{\rho}{\mu}\right)\right].$$

Again, removing the condition that N is given,

$$\text{Var } \hat{q}_k = \frac{(1-q_k)^2\mu_k^2}{E(D_k)}$$

where

$$E(D_k) = \left[\rho_k(1-q_k) + q_k\left(1 - \frac{\rho_k}{\mu_k}\right)\right] M \prod_{i=0}^{k-1} (1-\rho_i)(1-q_i). \tag{8}$$

Since $\hat{q}_L$ is biased under the alternative survivorship function, the measure of efficiency which has been chosen is

$$E_N = \frac{N \text{ Var } \hat{q}}{N \text{ Var } \hat{q}_L + NB^2} = \frac{N \text{ Var } \hat{q}}{N[MSE\ \hat{q}_L]}$$

where B, the asymptotic bias of $\hat{q}_L$, is $q_L^* - q$, and q_L^* is the result of replacing D and L in (4) by their expected values. Note that, since the bias of $\hat{q}_L$ is independent of N, the efficiency, for fixed q and ρ, is a decreasing function of N.

For a sub-interval on which $\rho = \frac{1}{3}$, $\mu = .2$, $q = .1813$, an interval "longer" in terms of ρ and q than would ordinarily be used, the efficiency[7] of $\hat{q}_L$ relative to $\hat{q}$, for $N = 1000$ is

$$E_{1000} = \frac{.189796}{.197244 + 1000(.00136)^2} = .95.$$

If the interval is shortened to $\mu = .1$ and $\rho = \frac{1}{6}$, the corresponding efficiency, for $N = 1000$, is .99.

Let the estimate of $P(T)$ based on the use of the maximum likelihood estimate (7) on each sub-interval be $\hat{P}(\hat{q}, T)$, and that based on the quadratic frequency estimate (4) on each sub-interval be $\hat{P}(\hat{q}_L, T)$. The efficiency measure is taken as the ratio of the variance of $\hat{P}(\hat{q})$ to the MSE of $\hat{P}(\hat{q}_L)$. All variances and biases are, again, asymptotic. The variances of the $\hat{q}_k$ are given by (8). The asymptotic variances of the $\hat{q}_{L_k}$ and of the two estimates of P(T) have all been obtained by use of the "delta method" [6, pp. 353–66].

The bias of

$$\hat{P}(\hat{q}_L, T) = \prod_{k=0}^{a-1} (1 - \hat{q}_{L_k})$$

[7] For variance $\hat{q}_L$ under the alternative see Appendix Note 3.

is taken as

$$B[\hat{P}(\hat{q}_L, T)] = \left[\prod_{k=1}^{a-1}(1 - \bar{q}_{L_k})\right] - P(T)$$

where each $\bar{q}_{L_k}$ is the result of replacing N, D and L in the estimate as given in (4) by the expectations of the appropriate N_k, D_k and L_k.

The expression for the asymptotic variance of $\hat{q}_{L_k}$, explicit in ρ and μ, is lengthy and of no practical value from a computing standpoint. The values of the Var $\hat{q}_{L_k}$ have been evaluated numerically, step by step, as described in Note 3 of the Appendix.

The "delta method," as applied to an estimate

$$\hat{P}(T) = \prod_{i=0}^{a-1}(1 - \hat{q}_i), \text{ where } \hat{q}_i \text{ is an estimate of } q_i,$$

gives,[8] for zero values of the asymptotic covariances between $\hat{q}_i$ and $\hat{q}_j$,

$$\text{Var } \hat{P}(T) = [P(T)]^2 \sum_{i=0}^{a-1} \frac{\text{Var } \hat{q}_i}{(1 - q_i)^2} \tag{9}$$

which is known as Greenwood's approximation [11]. This expression has been used for the asymptotic variance of each of the two estimates of $P(T)$.

The increase in the variance of the actuarial estimate with increasing number, a, of sub-intervals observed [13] in the case in which, for each a, $\rho_{a-1} = 1$, does not occur here with $(0, T)$ restricted so that some part of the histories are complete over $(0, T)$. The interval $(.8T, T)$ corresponds, for $a = 5$, to the last interval, for $a = 10$ to the last two intervals, and for $a = 20$, to the last four intervals. The contribution to M Var $\hat{P}(\hat{q}_L, T)$ of this part of the total period decreases with a, from .1923 for $a = 5$ to .1901 for $a = 20$.

For each a, the bias of $\hat{q}_L$ is maximum on the last sub-interval $(a-1, a)$. The bias on the last interval decreases from .00625 (approximately $.02q$) for $a = 5$, to .00017 (approximately $.002q$) for $a = 20$.

It is of interest to note that, if the familiar estimate

$$\hat{q}_{f_k} = \frac{D_k}{N_k - \dfrac{L_k}{2}} \tag{10}$$

is used on each sub-interval, and $\hat{P}(T)$ is estimated as $\prod_{k=1}^{a-1}(1-\hat{q}_{f_k})$, that the bias of the resulting estimate of P(T), is, for each value of a, approximately double that of the estimate based on $\hat{q}_{L_k}$. The approximation, $\hat{q}_{f_k}$, is not consistent under the general linear model. Since $\hat{q}_L$, is based on the same set of frequency observations, and shows better performance, the approximation (10)

[8] Since $\hat{q}_L$ is biased under exponential survivorship, (9) becomes

$$\text{Var } \hat{P}(\hat{q}_L, T) = [\bar{P}(T)]^2 \sum_{k=1}^{a-1} \frac{\text{Var } \hat{q}_{L_k}}{(1 - \bar{q}_{L_k})^2}$$

where

$$\bar{P}(T) = \prod_{k=1}^{a-1}(1 - \bar{q}_{L_k}).$$

will not be discussed further, although for use with intervals corresponding to small values of both q and ρ it will fulfill the accuracy requirements in some applications. The performance of the estimate of $P(T)$ based on the use of $\hat{q}_L$ on the sub-intervals, relative to that based on use of the maximum likelihood estimate on each sub-interval, is summarized in Table 436.

The Estimation of $P(t)$ Under the Alternative Model

In the illustration of the performance of $\hat{q}_L$ under the alternative model, the net survivorship is taken to be $P(t)=e^{-\mu t}$ on $(0, T)$, and $\mu T=2.0$, so that $P(T)=.135335$. The estimates based on $\hat{q}_L$ are discussed for $a=5$, 10 and 20 sub-intervals. The ρ_k values[9] have been chosen to approximate situations met in application. For each value of a, $\rho_k=5/[a+5(a-k)]$ for $k=0, 1, \cdots, a-1$.

TABLE 435

VALUES OF THE PARAMETERS USED IN THE ILLUSTRATION

No. of intervals a	ρ_0	ρ_1	$\rho_2 \cdots \rho_{a-2}$	ρ_{a-1}	$q=1-e^{-2/a}$
5	1/6	1/5	1/4 $\cdots$ 1/3	1/2	.32968
10	1/12	1/11	1/10 $\cdots$ 1/4	1/3	.18127
20	1/24	1/23	1/22 $\cdots$ 1/6	1/5	.09516

In the illustration, for any given a, the q values do not change from one sub-interval to another. For each value of a, $\prod_{i=0}^{5}(1-\rho_i)=\frac{1}{6}$, and of the total of M patients, $M\frac{1}{6}P(T)=.0225\,M$ are expected to survive under observation to time T. This schedule involves fairly heavy withdrawal rates, particularly in the first few intervals. In practice, the entries of the last few years are not used so that for the first two to five years of follow-up there are no incomplete histories.

The bias of the estimate $\hat{P}(\hat{q}_L, T)$ due to non-linearity of the survivorship function $P(t)=e^{-\mu t}$, is essentially eliminated by division of the period $(0, T)$ into $a=20$ sub-intervals. The reduction in bias is accomplished without loss in efficiency.

CONCLUSIONS

The actuarial method is appropriate for a large and important class of follow-up studies of survivorship in chronic disease. The frequency estimate, $\hat{q}_L$, arising as the maximum likelihood frequency estimate from the approximation of $P(t)$ by a polygonal function is a very satisfactory estimate of the interval probability. The clerical and computational work involved in its use is simple. It is appropriate for use with either the common closing date or the date of last report method of recording the sample information. The results presented are

[9] Note that the schedule of withdrawals corresponding to the sets of ρ values used here are those which would result from use of a common closing date applied to uniform entries, for which $R/M=\frac{1}{2}$ of the histories are complete to T, and $(M-R)/a=w$ histories become incomplete on each sub-interval. Then $R+(a-k)w$ histories are complete to k, and of these a proportion $\rho_k=w/[R+(a-k)w]$ become incomplete on k, $k+1$. While the ρk values chosen for the numerical example correspond to uniform entries, this is not a requirement under the general model, since the ρ_k are not subject to this restriction.

TABLE 436

VARIANCE, BIAS AND EFFICIENCY OF $\hat{P}(\hat{q}_L) = \prod_{k=0}^{a-1} (1-\hat{q}_{L_k})$ AS AN ESTIMATE OF $P(T)$

Number of sub-intervals	$a=5$	$a=10$	$a=20$
Sub-interval*	$q=.32968$	$q=.18127$	$q=.09516$
M Var $\hat{P}(\hat{q}, T)$	.30298	.32005	.32488
M Var $\hat{P}(\hat{q}_L, T)$	.33117	.32812	.32699
$B[\hat{P}(\hat{q}_L, T)]$	.00322	.00081	.00009
$E_{1000}[\hat{P}(\hat{q}_L, T)]$	.887	.973	.993
$E_{3000}[\hat{P}(\hat{q}_L, T)]$	.836	.970	.993

* For each a, the value of q is the same for all sub-intervals. The ρ_k values for each a, and for $k=0, 1, \cdots, a-1$ are given in Table 1LE

all based on large sample approximations and on asymptotic variances. Suitable restrictions must be placed on M, the total number of histories, and on the number of histories complete over the entire period $(0, T)$.

The conditions necessary for satisfactory results with the actuarial estimate $\hat{P}(\hat{q}_L)$ are simple.

1. The actuarial method, designed for the case in which the form of the survivorship function is unknown, does not lend itself to extrapolation. The period $(0, T)$ over which $P(t)$ is to be estimated must be restricted so that it is covered by a reasonable amount of experience under observation. As a practical rule, it might be recommended that T be chosen so that at least 25 persons be observed to survive the interval $(0, T)$.
2. Since approximation to $P(t)$ by a polygonal function is involved, the sub-intervals must not be too "long" in terms of ρ and q.

If the first of these conditions is met the bias may be controlled by increasing the number of intervals with a resulting gain in efficiency. Maximum interval lengths corresponding to $q=.10$ and $\rho=\frac{1}{3}$ might be recommended for general use, although in some studies maximum values of $q=.2$ and $\rho=\frac{1}{3}$ will give the accuracy required.

APPENDIX

Note 1. An Example

The first four columns of the table below are taken from those given by Parker, Dry, Willius and Gage [16, page 97, Table 3], and are based on dates of last report.

The authors base their estimates of $P(t)$ on the interval estimates (10)

$$\hat{q}_{f_k} = \frac{D_k}{N_k - \frac{L_k}{2}}$$

which gives a fifteen year survivorship of .143, whereas the use of $\hat{q}_L$ on each

interval leads to an estimate of .141. The intervals of one calendar year used are rather "long" in terms of the q_k and ρ_k.

SURVIVORSHIP IN ANGINA PECTORIS

1	2	3	4	5	6	7	8
		Last Report					
x	N	D	L	$\hat{q}_L$	$\hat{P}(\hat{q}_L, x)$	$\hat{\rho}$	$100\,\dfrac{\hat{SE}_{\hat{P}(x)}}{P(x)}$
0	2418	456	0	.18859	1.0000	0	
1	1962	226	39	.11642	.8114	.0211	1.0
2	1697	152	22	.09018	.7170	.0136	1.3
3	1523	171	23	.11318	.6523	.0160	1.5
4	1329	135	24	.10256	.5784	.0190	1.8
5	1170	125	107	.11228	.5191	.0969	2.0
6	938	83	133	.09560	.4608	.1489	2.2
7	722	74	102	.11078	.4168	.1496	2.5
8	546	51	68	.09996	.3706	.1311	2.9
9	427	42	64	.10682	.3336	.1583	3.2
10	321	43	45	.14491	.2979	.1511	3.7
11	233	34	53	.16659	.2548	.2481	4.4
12	146	18	33	.14035	.2123	.2431	5.4
13	95	9	27	.11152	.1825	.3010	6.5
14	59	6	23	.12845	.1622	.4166	7.6
15	30				.1413		9.5

Column
1. x = Years following diagnosis at clinic.
2. N = Number of persons alive at beginning of interval.
3. D = Number of deaths observed in the interval $(x, x+1)$.
4. L = Number of withdrawals alive in the interval $(x, x+1)$.
5. $\hat{q}_L$ from equation (4).
6. $\hat{P}(\hat{q}_L, x) = \prod_{x=1}^{x-1}(1-\hat{q}_{L_k})$; and 7. $\hat{\rho}$ from equation (5).
8. Var q_L evaluated from equation (6) and $\hat{SE}_{\hat{P}(x)}$ from Greenwood's approximation, (9).

Note 2. Other Frequency Estimates

Berkson [4] arrives at equation (1) heuristically by "considering q to be given as the observed deaths plus the deaths expected among the living withdrawals after withdrawal, divided by the total number living at the beginning of the interval," or by setting the estimate, for linear survivorship as

$$q = \frac{1}{N}\left\{D + \sum^{L} [\text{Prob (death in } T_m, 1) \mid \text{survival to } T_m]\right\} \qquad (1A)$$

which leads to the likelihood equation (1).

Littell [13] arrives at his "effective persons exposed" estimate, for exponential survivorship as the solution of

$$q = \frac{1}{N}\left\{D + E_T \sum^{W} [\text{Prob (death in } T_i, 1)]\right\} \qquad (2A)$$

where E_T denotes expectation with respect to T which is uniform over the

interval. Under linear survivorship (2A) gives $\hat{q}_W = D/(N - W/2)$, and, if expectation with respect to T is dropped, gives

$$\bar{q}_W = D \Big/ \left[N - \sum^{W} (1 - T_j) \right].$$

These estimates are available only if a common closing date is imposed.

It was surprising to the writer to find a survivorship function for which the maximum likelihood estimate of q involves only N, D, W and the $T_1, \cdots, T_W$.

Let

$$P(t) = \frac{1 - q}{1 - q + qt} \qquad \text{for } 0 \leq t \leq 1. \tag{3A}$$

Then

$$\phi = q^{D_1}(1 - q)^{S} \left[\prod^{L} \frac{1 - q}{1 - q + qT_m} \right] \left[\prod^{D_2} \frac{qT_n}{1 - q + qT_n} \right]$$

and

$$\frac{\partial \log \phi}{\partial q} = \frac{D}{q} - \frac{N - D}{1 - q} + \sum^{W} \left[\frac{1 - T_j}{1 - q + qT_j} \right] = 0$$

which is equivalent to the expression resulting from dropping the expectation with respect to T in the second term on the right of (2A), or to

$$q = \frac{1}{N} \left[D + q \sum^{W} \left(\frac{(1 - q)(1 - T_j)}{1 - q + qT_j} \right) \right].$$

It is interesting to note that, for this survivorship function, (1A) gives

$$\hat{q} = \frac{D}{N - \sum^{L} (1 - T_m)}. \tag{4A}$$

If expectation with respect to the T distribution were introduced in the second term on the right of (1A), the estimate of (4A) would reduce to the familiar form (10)

$$\hat{q}_{I_k} = \frac{D_k}{N - \dfrac{L_k}{2}}. \tag{5A}$$

The estimate $\hat{q}_{I_k}$ of (5A) has been widely used [2, 3, 16, 18] as an approximation to $\hat{q}_L$, which is the solution of the equation

$$\hat{q}_L = \frac{D}{N_k - \dfrac{L}{2 - \hat{q}_L}}.$$

Note that $\hat{q}_{I_k}$ is consistent under the survivorship function (3A).

The survivorship function of (3A) corresponds to a decreasing intensity of risk

$$r_D(t) = \frac{q}{1 - q + qt}$$

and is an interesting addition to $P(t) = e^{-\mu t}$ and $P(t) = 1 - qt$ which have constant and increasing intensities respectively.

Note 3. The Asymptotic Variance of the Estimate $\hat{q}_L$ Under Exponential Survivorship and Linear Withdrawals

On the typical interval, for N given, and $N = D + L + S$, $\hat{q}_L$ is a function of the random variables D and L.

Let

$$\hat{q}_L = f(D, L) = \frac{1}{2N}\left[2N - L + D - \sqrt{(2N - L + D)^2 - 8DN}\right] \quad (6A)$$

and let $f[E(D), E(L)]$ denote the result of replacing D and L by their expected values $E(D)$ and $E(L)$, in the right hand side of (6A). Expansion to linear terms, of $f(D, L)$ about $E(D)$, $E(L)$ gives

$$f(D, L) - f[E(D), E(L)] \cong [D - E(D)]f_D^* + [L - E(L)]f_L^* \quad (7A)$$

where f_D^* and f_L^* are the respective partial derivatives of $f(D, L)$ evaluated at $E(D)$, $E(L)$. Squaring and taking expectations on both sides of (7A)

$$E\{f(D, L) - f[E(D), E(L)]\}^2 \cong (f_D^*)^2 \operatorname{Var} D + (f_L^*)^2 \operatorname{Var} L + 2f_D^* f_L^* \operatorname{Cov}(D, L). \quad (8A)$$

For $\hat{q}_L$, and for D and L observations from a multinomial, the right hand side of (8A) may be taken [6, pp. 353–66] as the asymptotic variance of $\hat{q}_L$.

For

$$q^* = f[E(D), E(L)] \quad \text{and} \quad \Delta^{*2} = [2N - E(L) + E(D)]^2 - 8NE(D)$$

$$f_D^* = (2 - q^*)/\Delta^* \quad \text{and} \quad f_L^* = q^*/\Delta^*$$

and

$$\operatorname{Var} \hat{q}_L = \frac{1}{\Delta^{*2}}\{(2 - q^*)^2 \operatorname{Var} D + (q^*)^2 \operatorname{Var} L + 2q^*(2 - q^*) \operatorname{Cov}(DL)\}.$$

In the computations throughout N has been replaced by the appropriate expectation of N_k, that is by

$$E(N_k) = M \prod^{k-1} (1 - \rho_i)(1 - q_i).$$

The resulting variances are identical with those given by considering each $\hat{q}_{L_k}$ as a function of three random variables D_k, L_k, and $G_k = M - N_k$ having expectations

$$E(G_k) = M - M\prod^{k-1}(1-\rho_i)(1-q_i) = M - E(N_k)$$

$$E(D_k) = \left[q_k\left(1-\frac{\rho_k}{\mu_k}\right) + \rho_k(1-q_k)\right]E(N_k) \text{ and} \tag{9A}$$

$$E(L_k) = \frac{\rho_k q_k}{\mu_k}E(N_k).$$

Note that $M = G_k + D_k + L_k + N_{k+1}$

BIBLIOGRAPHY

[1] Berkson, J., Walters, W., Gray, H. K., and Priestly, J. T., "Mortality and survival in cancer of the stomach: A statistical summary of the experience of the Mayo Clinic," *Proceedings of the Staff Meetings of the Mayo Clinic*, 27 (1952), 137–51.

[2] Berkson, J., and Gage, R. P., "Survival curve for cancer patients following treatment," *Journal of the American Statistical Association*, 47 (1952), 501–15.

[3] Berkson, J., and Gage, R. P., "Calculation of survival rates for cancer," *Proceedings of the Staff Meetings of the Mayo Clinic*, 25 (1950), 270–86.

[4] Berkson, J., "Estimation of the interval rate in actuarial calculations," (summary) *Journal of the American Statistical Association*, 49 (1954), 363.

[5] Boag, J. W., "Maximum likelihood estimates of the proportion of patients cured by cancer therapy," *Journal of the Royal Statistical Society*, 11 (1949), 15–53.

[6] Cramér, Harald, *Mathematical Methods of Statistics*, Princeton, N. J., Princeton University Press, 1946.

[7] Densen, P. M., "Long-time follow-up in morbidity studies: The definition of the group to be followed," *Human Biology*, 22 (1950), 233–7.

[8] Dorn, H. F., "Methods of analysis for follow-up studies," *Human Biology*, 22 (1950), 238–48.

[9] Doull, J. A., Guinto, R. S., Rodriguez, J. N., and Bancroft, H., "The incidence of leprosy in Cordova and Talisay, Cebu, P.I.," *International Journal of Leprosy*, 10 (1942), 107–31.

[10] Fix, E., and Neyman, J., "A simple stochastic model of recovery, relapse, death and loss of patients," *Human Biology*, *23* (1951), 205–41.

[11] Greenwood, M., "The natural duration of cancer," *Reports on Public Health and Medical Subjects*, No. 33, 1926.

[12] Harris, T. E., Meier, P., and Tukey, J., "Timing of the distribution of events between observations: A contribution to the theory of follow-up studies," *Human Biology*, 22 (1950), 249–70.

[13] Littell, A. S., "Estimation of the *T*-year survival rate from follow-up studies over a limited period of time," *Human Biology*, 24 (1952), 87–116.

[14] Meier, P., "Non-parametric estimation of survivorship," (summary) *Journal of the American Statistical Association*, 50 (1955), 589.

[15] Neyman, Jersey, *First Course in Probability and Statistics*, New York, Henry Holt and Company, 1950.

[16] Parker, R. L., Dry, T. J., Willius, F. A., and Gage, R. P., "Life expectancy in angina pectoris," *Journal of the American Medical Association*, 131 (1946) 95–100.

[17] Seal, H. L., "The estimation of mortality and other decremental problems," *Skandinavisk Aktuarietidskrift* (1954) 137–62.

[18] Sloan, W. P., Bargen, J. A., and Gage, R. P., "Life histories of patients with chronic ulcerative colitis: A review of 2,000 cases," *Gastroenterology*, 16 (1950), 25–38.

GRAPHIC COMPUTATION OF TAU AS A COEFFICIENT OF DISARRAY

HAROLD D. GRIFFIN
Lincoln, Nebraska

1. INTRODUCTION 441
2. HOLMES' GRAPHIC COEFFICIENT OF CORRELATION 442
3. GRAPHIC COMPUTATION APPLIED TO A SIMPLE RANK-ORDER PROBLEM 442
3.1. Basic Procedure 442
3.2. Modification when Correlation is Patently Negative 443
4. TIED RANKS 443
4.1. Ties in One Ranking Only 443
4.2. Ties in Both Rankings 445
4.3 Proof that the Graphic Method for Ties gives Kendall's S 446
5. TABLE FOR $\binom{n}{2}$, AND FOR $\binom{t}{2}$ AND $\binom{u}{2}$ IN CALCULATING T AND U WITH TIED RANKS 447
6. REFERENCES 447

1. INTRODUCTION

THE real or fancied difficulties in computing M. G. Kendall's *tau* coefficient [1] have retarded its general use as a measure of correspondence between rankings. However, when *tau* is considered as a coefficient of disarray, that is, as a measure resultiug from the least number of moves required for transforming one ranking into the other, it will yield to simple graphic interpretation and computation. The formula for *tau* as a coefficient of disarray is

$$\tau = 1 - \frac{2s}{\binom{n}{2}}, \tag{2}$$

which may also be written as

$$\tau = 1 - \frac{4s}{n(n-1)}, \tag{2}$$

where

s = the minimum number of interchanges which will transform one of the sample rankings into the other, and

$$\binom{n}{2} = \frac{1}{2}\, n(n-1)$$

2. HOLMES' GRAPHIC COEFFICIENT OF CORRELATION

Back in the 1920's, a Canadian educator, S. D. Holmes, now text-book editor, Ontario Provincial Department of Education, Toronto, developed a graphic method for demonstrating correlation [2]. Although Mr. Holmes' purpose was merely to use this graphic method as an illustrative teaching device, he carried it through to the development of the formula

$$r_g = 1 - \frac{4X}{n(n-1)}, \tag{3}$$

where

r_g = the graphic coefficient of correlation, and

X = the number of intersections resulting from connecting the two rankings of each member by lines, one ranking having been put in natural order.

Note the similarity in structure of formulas (2) and (3).

Neither S. D. Holmes nor his mentor, Peter Sandiford, were aware that a formula for rank correlation had been uncovered. They were interested in r_g as an estimator of the Pearson product-moment correlation coefficient. When during the years 1929 to 1931 I sent Sandiford the results of studies made of r_g as an rank-order correlation coefficient, he discouraged publication as being of little practical value. Our uncertainty of its standard error of estimate seemed particularly disturbing to him.

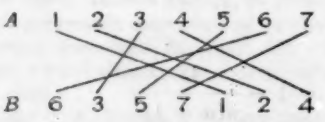
3. GRAPHIC COMPUTATION APPLIED TO A SIMPLE RANK-ORDER PROBLEM

3.1. *Basic Procedure*

Let us apply Holmes' graphic method to an illustrative problem given by Kendall [1], p. 7f. Consider the two rankings of seven:

A	1	2	3	4	5	6	7	
B	6	3	5	7	1	2	4	(Problem 1)

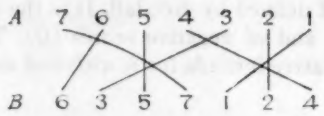
When writing this down we should leave enough space between the lines for satisfactory revelation of the intersections we shall make. Thus,



Connect similar ranks with lines, which need not be perfectly straight, counting intersections as the lines are drawn. We obtain 13 intersections, which, substituted in formula (3), gives $r_g = -0.238$. This agrees with the result obtainable by the traditional methods of computation. It is apparent that Holmes' r_g is the same as Kendall's τ, and that the graphic method can be used in obtaining *tau*. Consequently, formula (1) may be used for the graphic computation of *tau*, letting s be the symbol for the number of intersections.

3.2. *Modification when Correlation is Patently Negative*

If one set of ranks be reversed while the other is fixed, the value of *tau* changes in, and only in, sign. This reversal may be done to simplify computations when preliminary inspection indicates the value of *tau* will be negative. The array in Problem 1 makes it apparent that correlation is negative. In such cases as this, some labor can be eliminated by reversing the order of one line. Thus,



When one line or axis is reversed, change the sign of the value of the *tau* obtained.

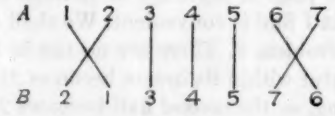
4. TIED RANKS

4.1. *Ties in One Ranking Only*

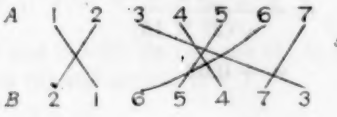
In case of tied ranks, the graphic method presents few difficulties. In approaching the problem of tied ranks graphically, we might assume in each case of ties that differences exist, but the measuring device failed to distinguish these differences. A literal acceptance of this statement might suggest that we construct a series of charts for all possible ways of assigning integral ranks to the tied measures, compute *tau* for each, and take the arithmetic mean of these values as the average coefficient. If there were merely one pair of tied ranks, this could be done, for there are only two possible arrangements for a pair of ties. From here on, difficulties are compounded, since there are six possible arrangements for three objects, 24 for four, 120 for five, and 720 for six! However, in practice, if the ties, irrespective of complexity, are in one ranking only, we need construct but two graphs, each eliminating the tied ranks—one minimizing, the other maximizing possible differences. Then we total the interscections from the two graphs, thus obtaining $2s$ to use in formula (1). An example will make the process clear. Let us consider these two rankings of seven:

A	1	2	3	4	5	6	7	
B	2	1	4½	4½	4½	7	4½	(Problem 2)

When we minimize differences, the rankings become



If we maximize differences, we obtain



From the above we get $2s = 10$, that is, $s_1 + s_2$; and $\tau_a = +0.524$.

If one seeks the refinement of Kendall's adjustment for tied ranks [1], Chapter 3, the graphic method supplies the needed information, since

$$S = \binom{n}{2} - 2s, \tag{4}$$

where S is a symbol used and defined by Kendall. It is the summation of the totals of positive scores (P) and of negative scores (Q). This provides the numerator in Kendall's alternative formula for τ_b with tied ranks [1], p. 35, 2d ed., which he writes as

$$\tau_b = \frac{S}{\sqrt{\{\frac{1}{2}n(n-1) - T\}}\sqrt{\{\frac{1}{2}n(n-1) - U\}}}, \tag{5}$$

where

$T = \frac{1}{2}\sum_t t(t-1)$ or one-half the summation of the various sets of ties in one ranking, and

$U = \frac{1}{2}\sum_u u(u-1)$ or one-half the summation of the various sets of ties in the other ranking.

Personally, I prefer a restatement of formula (5) as

$$\tau_b = \frac{S}{\sqrt{\left\{\binom{n}{2} - \Sigma\binom{t}{2}\right\}}\sqrt{\left\{\binom{n}{2} - \Sigma\binom{u}{2}\right\}}}, \tag{6}$$

which makes it possible to substitute values of $\binom{n}{2}$, $\binom{t}{2}$, and $\binom{u}{2}$ for all elements in the denominator of formula (6) directly from section 5 of this paper. This "Table for $\binom{n}{2}$, and for $\binom{t}{2}$ and $\binom{u}{2}$ in Calculating T and U with Tied Ranks" is one which I wrote on an end page of my copy of Kendall's *Rank Correlation Methods*. It is elementary, but I find it convenient. We shall apply formula (6) and data from section 5 to Problem 2. There are no ties in the A-ranking, so the first half of the denominator within its braces becomes $21 - 0 = 21$. There is one 4-fold tie in the B-ranking, so the second half becomes $21 - 6 = 15$. Hence,

$$\tau_b = \frac{11}{\sqrt{(21 \times 15)}}$$
$$= +0.62$$

4.2. *Ties in Both Rankings*

Where ties are found in both rankings, it is best to construct four graphs—minimizing and maximizing the differences for each ranking. Consider these two heavily-tied rankings of ten:

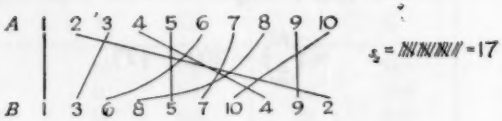
A	2	2	2	4.5	4.5	8	8	8	8	8	
B	5	2.5	1	5	7.5	2.5	9	5	10	7.5	(Problem 3)

Minimizing both, we obtain

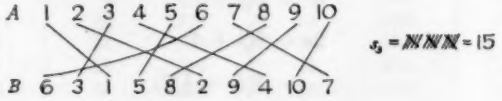
Maximizing both,

A	3	2	1	5	4	10	9	8	7	6
B	6	3	1	5	8	2	9	4	10	7

but note: In all methods for obtaining *tau* graphically, it is vital that one of the two rankings be in natural order. Therefore, as both rankings have been disarranged in maximizing differences, it will be necessary to make suitable adjustment before drawing our lines. Thus:



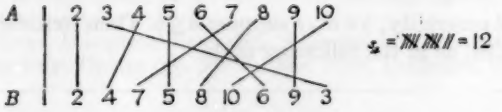
Minimizing *A*, maximizing *B*:



Maximizing *A*, minimizing *B*:

A	3	2	1	5	4	10	9	8	7	6
B	4	2	1	5	7	3	9	6	10	8

Again, a rearrangement is necessary:



Therefore, $4s = 54$ and $2s = 27$. By formula (1), $\tau_a = +0.40$, and by formula (4), $S = 18$. By formula (6) and section 5,

$$r_b = \frac{18}{\sqrt{\{45 - (3 + 1 + 10)\}}\sqrt{\{45 - (3 + 1 + 1)\}}}$$

$$= \frac{18}{\sqrt{(31 \times 40)}}$$

$$= + 0.511$$

Construction of more than four graphs in solving any rank-correlation problem involving two variables appears to be unjustified.

4.3 *Proof that the Graphic Method for Ties gives Kendall's S*[1]

A simple proof that the graphic method for ties will yield Kendall's S may be had as follows: Let T be the number of pairs of corresponding ranks for which at least one tie appears. Then

$$\binom{n}{2} = P + Q + T,$$

and

$$S = P - Q$$

$$= \binom{n}{2} - 2(Q + \tfrac{1}{2}T)$$

Also,

$$4s = 4Q + 2T,$$

$$2s = 2Q + T$$

$$= 2(Q + \tfrac{1}{2}T).$$

That $4s = 4Q + 2T$ may be seen as follows: consider each pair of corresponding rankings

$$\begin{Bmatrix} a \\ c \end{Bmatrix}, \begin{Bmatrix} b \\ d \end{Bmatrix}$$

Without loss of generality, we may suppose $a \leq b$. Then consider the six possible sign relationships, as in the following table:

---

[1] For the proof I am indebted to William H. Kruskal, who has also contributed a number of other improvements to the paper.

		Contributions to					
$a \lessgtr b$	$c \gtrless d$	Q	T	Min A– Min B s	Min A– Max B s	Max A– Min B s	Max A– Max B s
<	<	0	0	0	0	0	0
<	=	0	1	0	1	0	1
<	>	1	0	1	1	1	1
=	<	0	1	0	0	1	1
=	=	0	1	0	1	1	0
=	>	0	1	1	1	0	0

But in each row the contribution to $4Q+2T$ is the sum of the contributions to the four s's. (Actually only the first three rows of the table are really needed.) Therefore,

$$S = \binom{n}{2} - 2s.$$

This proof provides at the same time formal proofs for the simpler cases, with ties in only one or in neither ranking.

5. TABLE FOR $\binom{n}{2}$, AND FOR $\binom{t}{2}$ AND $\binom{u}{2}$ IN CALCULATING T AND U WITH TIED RANKS

n	$\binom{n}{2}$	n	$\binom{n}{2}$	n	$\binom{n}{2}$	n	$\binom{n}{2}$
2	1	15	105	27	351	39	741
3	3	16	120	28	378	40	780
4	6	17	136	29	406	41	820
5	10	18	153	30	435	42	861
6	15	19	171	31	465	43	903
7	21	20	190	32	496	44	946
8	28	21	210	33	528	45	990
9	36	22	231	34	561	46	1035
10	45	23	253	35	595	47	1081
11	55	24	276	36	630	48	1128
12	66	25	300	37	666	49	1176
13	78	26	325	38	703	50	1225
14	91						

6. REFERENCES

[1] Kendall, Maruice G., *Rank Correlation Methods*, London, Charles Griffin & Company Limited. 1st Edition (1948), 2d Edition (1955). American agents, New York, Hafner Publishing Company.

[2] Sandiford, Peter, *Educational Psychology*, Appendix B, A Graphical Method of Estimating R for Small Groups (pp. 391–4), New York, Longmans, Green & Co. (1928).

A NEW BIVARIATE SIGN TEST

ISADORE BLUMEN
Cornell University

A sign test for the hypothesis that the medians of two or more variables (not necessarily independent) have a particular value is described in detail for the bivariate case. The slopes of the vectors from the bivariate median to the n sample points are ordered and we let $a_j = \pm 1$ according as the vector corresponding to the jth slope is above or below a horizontal line through the median. The bivariate test is based on the statistic $v^2 = 2(v_1^2 + v_2^2)/n$, where $v_1 = \sum a_j \cos(\pi j/n)$, $v_2 = \sum a_j \sin(\pi j/n)$, and the sum is over the integers $j = 1, 2, \cdots, n$. The statistic has the chi-square distribution for large samples. The power of this test relative to other possible sign tests and Hotelling's T^2 is discussed.

1. BACKGROUND

LET us suppose that there are two measurements of an individual's behavior in a particular situation. In a factory, for example, one measurement might be an index of the quality of the individual's output while the other might be the quantity produced. Fifteen machines are chosen at random and the first and second shift workers are to be compared.

If we compare the workers on only one variate, say quality, the problem is a standard one [2]. For if there is no effect of shift and if the workers have been chosen at random from some relevant population, the quality of the output recorded on the first shift and the quality recorded on the second shift could be viewed as independent chance variables which have the property that the median of the differences is zero. That is, if we write a plus whenever quality is better on the first shift and a minus whenever it is better on the second shift, the fifteen signs can be taken as coming from a binomial distribution with $n = 15$ and $p = \frac{1}{2}$.

If quality on one shift were better than that on the other, the median of the differences would be different from zero and the signs would follow a binomial distribution with $p \neq \frac{1}{2}$. Hence, a reasonable test procedure for the hypothesis that there is no shift effect is to reject the hypothesis whenever the number of plus or minus signs is excessive.

This test is the grandfather of most nonparametric procedures and shares their virtues and vices. From our point of view, the principal virtue is that it makes no requirement that the behavior of quality scores on different machines must be the same. In fact, it does not even require that the variability of quality scores for workers on a particular machine be the same, except that these operators have scores which vary about some common level for the machine. A machine whose operators produce a highly variable product will not, therefore, dominate this test as might happen if one used the raw differences. (Of course, to the extent that the differences actually are homogeneous the sign test may not make most efficient use of the data.)

If we are to compare the two shifts on both quantity and quality we face an added complication in the fact that these are related measures for which the relationship is not simple. This dependence makes it impossible to perform sign tests on quantity and quality separately and treat the tests as independent. An extreme example will make the point clear: Suppose that whenever there is a plus on quality differences there must be a plus on quantity differences and similarly for minuses. A decision to accept or reject a hypothesis on quality will then imply the same decision about quantity.

In practice, of course, the association is not this close and extra information is gained by looking at both variates. When the differences are from a common normal distribution, an appropriate statistic is Hotelling's T^2 [1].

In the event that the requirement of normality cannot be met by the data, a bivariate sign test may be appropriate. It is possible to devise many such tests and before describing the one which we recommend it may be worth looking at a few.

We begin by plotting the differences on a graph. The simplest test then is probably to divide the plane up by two lines through the origin. There will be four "quadrants" and the two pairs of opposite quadrants can be separately considered. There will be n_1 points falling into the first pair, n_2 into the second. Of the n_1, s_1 will fall into one of the quadrants and s_1 is approximately normally distributed under any reasonable restrictions on the way the lines are drawn. In particular, we can draw the lines so that $n_1 = n/2$ (for convenience, assume n even) and so that the angle containing the n_1 observations is as small as possible. Then it is not hard to show that

$$L^2 = \frac{4\left(s_1 - \frac{n_1}{2}\right)^2}{n_1} + \frac{4\left(s_2 - \frac{n_2}{2}\right)^2}{n_2} \tag{1.1}$$

has, approximately, a chi square distribution with two degrees of freedom under our null hypothesis and will tend to be large when the null hypothesis is false. The test then is to reject when L^2 is large. Unfortunately, although easy to apply, the test has much less possibility of picking up the fact that our null hypothesis is false than the V^2 test proposed later since we effectively use only half our observations for each possible direction of shift.

A test proposed by Hodges [3] is not much more difficult to apply. He suggests drawing lines through the origin and counting the number of observations falling on either side of each line. One looks for a line for which the smallest number fall on one side and bases the test on this number. Put crudely, one relies on the observations to find the combination of quantity and quality least favorable to supporting the hypothesis and then applies a sign test. Tables for Hodges' procedure for many reasonable combinations of sample size and significance level may be found in [3]. The theory for the procedure is still incomplete. A heuristic argument for the point of view that the procedure has relatively little power is given in the last section.

2. THE TEST PROCEDURE

The univariate sign test involves the reduction of the problem of testing the hypothesis that the median of a distribution has a particular value to a binomial probability problem. It does this by taking as the basic data for the statistical test only the directions in which the original observations differ from the hypothesized value. There are two possible directions, to which plus and minus signs are attached. Our bivariate generalization is based on the directions of the vectors from the hypothesized value to (x_i, y_i).

In the particular example we are using, the hypothesis is that the median of quality and quantity differences between shifts is zero. Thus, if x_i' is the measure of quality on the ith machine on the first shift and x_i'' is the measure on the second shift, then the difference $x_i = x_i'' - x_i'$ is viewed as an observation from some population for which the median is zero. Similarly, if we use y to measure quantity, the differences $y_i = y_i'' - y_i'$ have median zero. Generally, we have that the median of $x_i \cos\theta + y_i \sin\theta$ is zero for any θ.

The artificial data of the first two lines of the table below may be used to illustrate the procedure. The first machine on the first shift showed a quantity value of 20.4 and a quality value of 15.7. The corresponding values for the

Machine		1	2	3	4	5	6	7	8
First shift	x_i'	20.4	20.3	19.5	18.0	30.3	12.7	23.8	32.0
	y_i'	15.7	29.1	30.3	28.9	17.6	21.8	16.2	16.8
Second shift	x_i''	10.7	33.7	35.6	40.0	12.5	24.3	43.4	16.3
	y_i''	35.5	3.4	20.5	30.2	26.4	22.6	33.2	39.4
Differences	x_i	−9.7	13.4	16.1	22.0	−17.8	11.6	19.6	−15.7
	y_i	19.8	−25.7	−9.8	1.3	13.8	0.8	17.0	22.6
Ratios	y_i/x_i	−2.04	−1.92*	−.61*	.06	−.78	.07	.87	−1.44

Machine		9	10	11	12	13	14	15
First shift	x_i'	3.2	21.7	24.8	24.6	21.8	23.8	27.0
	y_i'	13.8	22.3	20.9	8.7	25.8	7.7	14.4
Second shift	x_i''	25.3	33.4	15.8	21.1	34.5	33.4	38.7
	y_i''	23.9	42.1	21.8	35.1	24.5	36.1	21.1
Differences	x_i	22.1	11.7	−9.0	−3.5	12.7	9.6	11.7
	y_i	10.1	19.8	0.9	26.4	−1.3	28.4	9.7
Ratios	y_i/x_i	.46	1.69	−.10	−7.54	−.102*	2.96	.83

* $y_i < 0$

second shift are 10.7 and 35.5. Hence, the difference in quantity is $x_1 = x_1'' - x_1' = -9.7$, while $y_1 = 19.8$ is the quality difference. Similar differences are computed for each machine, giving the fifth and sixth rows below. These are to be the basic data for the sign test, since the medians of these variates should be zero if there is no shift effect.

The direction of the vector to (x_i, y_i) from the origin (with respect to the positive x axis) is given if one knows the slope y_i/x_i and the sign of y_i. Any test based on the direction of the vectors can, therefore, be based on these quantities. Our last line gives the slopes for each pair of observations, with an asterisk to indicate a negative value of y_i.

The probability that y_i has an asterisk is one half, since under the null hypothesis the median of the conditional distribution on any line through the origin is zero, and fixing the ratio y_i/x_i simply fixes on a single line. Hence, if we look at the directions of the vectors with similar slopes, there should be no preferred direction since a particular vector and the one obtained by a rotation of 180° are equally likely. On the other hand, if the medians are different from zero, the vectors from the origin will have a preferred direction.

We can arrange the vectors according to angle with the positive horizontal axis, beginning with the smallest and going to the largest angle. For our data, this yields .06, .07, .46, .83, .87, 1.69, 2.96, −7.54, −2.04, −1.92*, −1.44, −.78, −.61*, −.102*, −.10.

We can standardize the problem by changing the slopes of adjacent vectors so that the angles the slopes represent will be $180/n$ degrees apart. This standardization will give us a test which does not depend on the unit of measurement used or on the correlation between the two variates.

The standardized vectors will tend to point in the same direction if the median is not zero and in a random pattern if it is zero. Put another way, if we take the intersection of the standardized vectors (prolonged if necessary) and the unit circle, then the center of gravity of these points should be near the origin if the null hypothesis is true and different from the origin otherwise.

This center of gravity is obtained by computing the quantities

$$v_1/n = \frac{1}{n}\sum_{j=0}^{n-1} a_j \cos\left(\frac{180j}{n}\right)^\circ$$

and (2.1)

$$v_2/n = \frac{1}{n}\sum_{j=0}^{n-1} a_j \sin\left(\frac{180j}{n}\right)^\circ$$

where a_j is plus or minus one according to whether or not there is an asterisk at the $(j+1)$st ordered slope. In the above example

$$\begin{aligned} v_1 &= \cos 0^\circ + \cos 12^\circ + \cos 24^\circ + \cos 36^\circ + \cos 48^\circ + \cos 60^\circ + \cos 72^\circ \\ &\quad + \cos 84^\circ + \cos 96^\circ - \cos 108^\circ + \cos 120^\circ + \cos 132^\circ - \cos 144^\circ \\ &\quad - \cos 156^\circ + \cos 168^\circ \\ &= 5.064 \end{aligned}$$

$$\begin{aligned} v_2 &= \sin 0^\circ + \sin 12^\circ + \sin 24^\circ + \sin 36^\circ + \sin 48^\circ + \sin 60^\circ + \sin 72^\circ \\ &\quad + \sin 84^\circ + \sin 96^\circ - \sin 108^\circ + \sin 120^\circ + \sin 132^\circ - \sin 144^\circ \\ &\quad - \sin 156^\circ + \sin 168^\circ \\ &= 5.622 \end{aligned}$$

In the next section it is shown that if the null hypothesis is true then v_1 and v_2 are, approximately, independent normal variables with mean zero and variance $n/2$. Under the alternative, the means of one or both are different from zero. It follows that we can use the chi square tables with two degrees of freedom to test our hypothesis, rejecting if

$$v^2 = 2(v_1^2 + v_2^2)/n \tag{2.2}$$

is greater than χ_α^2. By way of illustration, if we use the five per cent level in our example, we should reject since

$$v^2 = 2(5.06^2 + 5.62^2)/15$$
$$= 8.0 > 5.99.$$

3. THE THEORY OF THE TEST

The basic data for the test can be taken to be n independent pairs of observations $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ from a bivariate probability distribution for which there is assumed to exist a unique point (μ, ν) such that

$$\Pr\{a(X - \mu) + b(Y - \nu) > 0\} = 1/2 \tag{3.1}$$

for all real a and b. This point will be called the median. (Implicit in the above definition is a continuity condition on the distribution at (μ, ν). For ease of exposition, it will also be assumed that the bivariate distribution has a probability density. There is little difficulty in modifying definitions and procedures to suit cases where this is not true.) Like other bivariate sign tests this one is based on $u_1, u_2, \cdots, u_n$ only, where u_i is the angle which the vector from (μ, ν) to (x_i, y_i) makes with the vector determined by some arbitrary vector, say the one from (0, 0) to (1, 0).

With no loss in generality, we can concern ourselves with the test for the hypothesis that our median is zero. That is, choose between the hypotheses

$$H_0\colon \mu^2 + \nu^2 = 0 \tag{3.2}$$

and

$$H_A\colon \mu^2 + \nu^2 > 0. \tag{3.3}$$

It seems reasonable to consider only such sign tests as are invariant under H_0 with respect to the scale of measurement used and to rotation of the axes. In order to meet these requirements and to provide a reasonably simple test, proceed as follows:

(1) For each i, place a plus sign at the intersection between the unit circle and the vector (x_i, y_i) (prolonged if necessary) and a minus sign at the intersection between the unit circle and $(-x_i, -y_i)$.
(2) Beginning with any sign and proceeding in a given direction on the circle, say counter-clockwise, move the next sign forward or backward until its distance along the circle from the first is the arc for $180/n$ degrees.
(3) Proceed in this way until the signs are spaced an equal distance apart on the unit circle.

(4) Obtain the center of gravity of this system, attaching mass plus one to the points with positive signs and zero to the points with negative signs.

The test is based on the distance of this point from (0, 0). When H_0 is true this distance is clearly invariant under rotations and changes in either scale of measurement.

It is easy to obtain the limit distribution of our test statistic under H_0. Let

$$V_1 = \sum_{j=0}^{n-1} a_j \cos\left(\frac{\pi j}{n} + A\right)$$

and

$$V_2 = \sum_{j=0}^{n-1} a_j \sin\left(\frac{\pi j}{n} + A\right),$$

where the angle is measured from an arbitrary but fixed vector determined in advance and A is the angle from this vector to the first of the m points determined in step (3). Under H_0, the a_j are independent chance quantities which are plus one with probability one half and minus one with probability one half. The distance of the center of gravity in (4) from (0, 0) is

$$\sqrt{(V_1^2 + V_2^2)/n}.$$

Under H_0, the first two moments of V_1 are

$$E(V_1) = \sum_{j=0}^{n-1} \left\{\cos\left(\frac{\pi j}{n} + A\right)\right\} E(a_j)$$

$$= 0 \text{ since } E(a_j) = 0.$$

$$\text{Var } (V_1) = \sum_{j=0}^{n-1} \left\{\cos\left(\frac{\pi j}{n} + A\right)\right\}^2 \text{Var } (a_j).$$

Now, Var $(a_j) = 1$ and

$$\left(\cos\frac{\pi j}{n} + A\right)^2 = \left[\cos\left(\frac{2\pi j}{n} + 2A\right) + 1\right] \Big/ 2.$$

Hence,

$$\text{Var } (V_1) = \frac{n}{2} + \frac{1}{2}\sum_{j=0}^{n-1} \cos\left(\frac{2\pi j}{n} + 2A\right)$$

$$= \frac{n}{2}.$$

The distribution of $\sqrt{2/n}\, V_1$ may be approximated by the normal for sufficiently large n by virtue of the central limit theorem. That the theorem holds for our statistic follows from Liapounov's condition [1] on the second and third absolute moments. For this problem this condition is that

$$\lim_{n\to\infty} n^{-3/2} \sum_{j=0}^{n-1} E\left|a_j \cos\left(\frac{\pi j}{n} + A\right)\right|^3 = 0. \tag{3.4}$$

This is clearly true because

$$\left| a_j \cos\left(\frac{\pi j}{n} + A\right)\right|^3 \leqq 1.$$

It will be shown that the limit distribution of V_1 and V_2 is bivariate normal by proving that all linear combinations of V_1 and V_2 are normal. This follows easily from the fact that we can write

$$\sum_{j=0}^{n-1} \cos\left(\frac{\pi j}{n} + C\right) = \sum_{j=0}^{n-1}\left\{\cos\left(\frac{\pi j}{n} + A\right)\cos(C - A) - \sin\left(\frac{\pi j}{n} + A\right)\sin(C - A)\right\}. \tag{3.5}$$

That the left hand side has a normal limit follows from our discussion for V_1, since A was arbitrary. If we take $(C-A)$ to be $\pi/2$, it now follows that V_2 is normal. Furthermore, if we take $C-A$ so that $b_1/(b_1^2+b_2^2)^{1/2}=\cos(C-A)$ then $b_1V_1+b_2V_2$ is, except for a multiplicative constant, of the form given in (3.5) and also has a normal limit.

Since the variance of V_1 and V_2 is the same for any A, the limit distribution of $2/n$ (V_1, V_2) is bivariate normal with mean zero, variance unity, and zero covariance. The distribution of the sum of squares of two standardized independent normal variables is a chi square with two degrees of freedom and it follows that

$$V^2 = 2(V_1^2 + V_2^2)/n \tag{3.6}$$

has that distribution approximately.

An indication of how good the above approximation to the actual distribution is can be obtained by comparing the fourth moment of $\sqrt{2/n}\,V_1$ with the fourth moments of standardized normal and binomial variables. A computation very similar to that for Var (V_1) yields $(4/n^2)E(V_1^4)=3-3/n$, while the fourth moments for the normal and binomial $(p=\frac{1}{2})$ are 3 and $3-2/n$ respectively. The fourth moments of the distribution of $\sqrt{2/n}\,V_1$ and of the binomial will agree when there are 50% more observations for the bivariate problem. This suggests that the approximation to the V^2 distribution which is proposed here will be satisfactory for sample sizes about 50% larger than those for which the binomial is satisfactorily approximated by the normal.

The exact distribution for V^2 does not seem attainable in general except by enumeration of the 2^n possibilities corresponding to $\pm a_1, \pm a_2, \cdots, \pm a_n$ with some minor simplifications because of symmetry. An exception is the largest value of V^2 for each n whose probabilities are the same as those of the most extreme value of Hodges' statistic, i.e., $n2^{-n+1}$. It follows from this that the smallest sample sizes needed to get significance for nonrandomized tests at the .10, .05 and .01 levels are $n=8$, 9, and 12 respectively.

The relative power of the V^2 test for large samples is now obtained for the case where the alternatives are bivariate normal with means different from zero. Since there is no real restriction involved, we will assume independence

between the two variates and unit variance for each, and following Pitman [4], we take the mean of (X, Y) to be $(\Delta/\sqrt{n}, 0)$ where $\Delta>0$. In order to simplify the argument we will introduce a pair of new chance variables asymptotically equivalent to V_1 and V_2.

Let

$$W_1 = \sum_{i=1}^{n} \cos U_i \quad \text{and} \quad W_2 = \sum_{i=1}^{n} \sin U_i.$$

Then, since

$$E \cos U_i = \frac{1}{\sqrt{2\pi}} \frac{\Delta}{n} \int_0^{\pi} \cos^2 u e^{-\Delta^2(1+\cos^2 u)/2n}\, du$$

$$\sim \frac{\Delta}{\sqrt{2\pi n}} \int_0^{\pi} \cos^2 u\, du$$

$$= \frac{\Delta}{2} \sqrt{\pi/2n}$$

it follows that

$$E\left(\sqrt{\frac{2}{n}} W_1\right) \sim \frac{\Delta}{2} \sqrt{\pi}. \tag{3.6}$$

It is obvious that $E(W_2) = 0$.

The argument used to obtain the distribution of V_1 and V_2 under the null hypothesis can be applied here and shows that W_1 and W_2 are asymptotically independent normal variables with variance $n/2$. Hence for large but fixed n, $W^2 = 2(W_1^2 + W_2^2)/n$ has the non-central chi square distribution with parameter of non-centricity

$$\phi^2 = \frac{\pi\Delta^2}{4}$$

which can be compared with the corresponding value for Hotelling's T^2 which is $\phi^2 = \Delta^2$. That is, when the alternative is bivariate normal Hotelling's T^2 requires about 79% of the observations needed by this sign test to attain the same power.

As was pointed out earlier, the Hodges test seems to be less efficient. At the five per cent level, for example, the minimum number for significance is about $n/2 - 3\sqrt{n/4}$ for the range for which the distribution is known. If the search for the appropriate direction is at all efficient then the power will be approximately that of a sign test with 3 in place of 1.96.

But if the two variates are known to be independent, then a more powerful test can easily be constructed. Count the plus observations on one side of an arbitrary line. The difference of the count from $n/2$, divided by $\sqrt{n/4}$ is normal with zero mean and unit variance. Take a line at right angles, obtain the difference in count from $n/2$ and divide by $\sqrt{n/4}$. The sum of squares of these differences is approximately a chi square variable and the five per cent critical value

is, therefore, about six. This test is based on differences of order $\sqrt{n}$ from $n/2$, as is the Hodges test. But a difference of about $\sqrt{6n}$ from the origin is enough in this test as compared to about $3\sqrt{n}$ for the Hodges procedure. Put another way, the ratio of the number of observations needed to get the same power is about two to three.

The requirement of independence for the test just constructed was not necessary. Had we worked with the standardized vectors there would have been asymptotic independence of the orthogonal quantities. The relative power for this substitute test compared to T^2 is the same as that of the univariate sign test relative to Student's t, namely $2/\pi$.

REFERENCES

[1] Cramér, Harald, *Mathematical Methods of Statistics*, Princeton: Princeton University Press, 1946.

[2] Dixon, Wilfrid J., and Frank J. Massey, Jr., *Introduction to Statistical Analysis*, New York: McGraw-Hill Book Company, 1951.

[3] Hodges, J. L. Jr., "A bivariate sign test," *Annals of Mathematical Statistics*, 26, 3 (1955), 523–7.

[4] Noether, G. E., "On a theorem of Pitman," *Annals of Mathematical Statistics*, 26, 1 (1955), 64–8.

NONPARAMETRIC ESTIMATION FROM INCOMPLETE OBSERVATIONS*

E. L. KAPLAN
University of California Radiation Laboratory
AND
PAUL MEIER
University of Chicago

In lifetesting, medical follow-up, and other fields the observation of the time of occurrence of the event of interest (called a *death*) may be prevented for some of the items of the sample by the previous occurrence of some other event (called a *loss*). Losses may be either accidental or controlled, the latter resulting from a decision to terminate certain observations. In either case it is usually assumed in this paper that the lifetime (age at death) is independent of the potential loss time; in practice this assumption deserves careful scrutiny. Despite the resulting incompleteness of the data, it is desired to estimate the proportion $P(t)$ of items in the population whose lifetimes would exceed t (in the absence of such losses), without making any assumption about the form of the function $P(t)$. The observation for each item of a suitable initial event, marking the beginning of its lifetime, is presupposed.

For random samples of size N the product-limit (PL) estimate can be defined as follows: List and label the N observed lifetimes (whether to death or loss) in order of increasing magnitude, so that one has $0 \leq t_1' \leq t_2' \leq \cdots \leq t_N'$. Then $\hat{P}(t) = \prod_r [(N-r)/(N-r+1)]$, where r assumes those values for which $t_r' \leq t$ and for which t_r' measures the time to death. This estimate is the distribution, unrestricted as to form, which maximizes the likelihood of the observations.

Other estimates that are discussed are the actuarial estimates (which are also products, but with the number of factors usually reduced by grouping); and reduced-sample (RS) estimates, which require that losses not be accidental, so that the limits of observation (potential loss times) are known even for those items whose deaths are observed. When no losses occur at ages less than t, the estimate of $P(t)$ in all cases reduces to the usual binomial estimate, namely, the observed proportion of survivors.

CONTENTS

1. Introduction 458
 1.1 Formulation 458
 1.2 Nonparametric estimation 459
 1.3 Examples of the RS and PL estimates 459
 1.4 Notation 461
2. The Product-Limit Estimate 462
 2.1 Definition and calculation 462
 2.2 Mean and variance of $\hat{P}(t)$ 465
 2.3 Mean lifetime 467
3. The Reduced-Sample Estimate vs. the PL 469
 3.1 Alternatives to the PL estimate 469
 3.2 Dependence of deaths and losses 470

* Prepared while the authors were at Bell Telephone Laboratories and Johns Hopkins University respectively. The work was aided by a grant from the Office of Naval Research.

4. Actuarial Estimates 471
4.1 Estimates using only n, δ, λ 471
4.2 Estimates using average ages 473
5. Maximum Likelihood Derivation of the PL 475
6. Means and Variances 476
6.1 The PL estimate $\hat{P}(t)$ 476
6.2 Covariances and mean lifetimes 478
7. Consistency; Testing with Replacement 479

1. INTRODUCTION

1.1 *Formulation.* In many estimation problems it is inconvenient or impossible to make complete measurements on all members of a random sample. For example, in medical follow-up studies to determine the distribution of survival times after an operation, contact with some individuals will be lost before their death, and others will die from causes it is desired to exclude from consideration. Similarly, observation of the life of a vacuum tube may be ended by breakage of the tube, or a need to use the test facilities for other purposes. In both examples, incomplete observations may also result from a need to get out a report within a reasonable time.

The type of estimate studied here can be briefly indicated as follows. When a random sample of N values, $T_1, T_2, \cdots, T_N$ of a random variable is given, the sample distribution function $\hat{F}(t)$ is naturally defined as that which assigns a probability of $1/N$ to each of the given values, so that $\hat{F}(t)$ equals $1/N$ times the number of sample values less than the argument t. Besides describing the sample, this $\hat{F}(t)$ is also a nonparametric estimate of the population distribution, in the sense indicated in 1.2 below. When the observations are incomplete, the corresponding estimate is still a step-function with discontinuities at the ages of observed deaths, but it can no longer be obtained as a mere description of the sample.

The samples considered in this paper are incomplete in the sense that one has given, not a random sample $T_1, \cdots, T_N$ of values of the random variable T itself (called the lifetime), but the *observed lifetimes*

$$t_i = \min (T_i, L_i), \qquad i = 1, 2, \cdots, N. \tag{1a}$$

Here the L_i, called *limits of observation*, are constants or values of other random variables, which are assumed to be independent of the T_i unless otherwise stated (in Sections 3.2 and 7). For each item it is known whether one has

$$T_i \leqq L_i, \qquad t_i = T_i \text{ (a } \textit{death}\text{)} \tag{1b}$$

or

$$L_i < T_i, \qquad t_i = L_i \text{ (a } \textit{loss}\text{)}.$$

Ordinarily the T_i and L_i are so defined as to be necessarily nonnegative.

The items in the sample are thus divided into two mutually exclusive classes, namely deaths and losses. A loss by definition always precludes the desired knowledge of T_i. On the other hand, a death does not always preclude the knowledge of the corresponding L_i, in case the limits of observation are nonrandom and foreseeable. Such knowledge of the L_i may have value; for example,

it makes available the reduced-sample estimate (Section 3), if one chooses to use it.

The type of sample described is a generalization of the censored sample defined by Hald [17], and a specialization of the situation considered by Harris, Meier, and Tukey [18].

The term death has been adopted as being at least metaphorically appropriate in many applications, but it can represent any event susceptible of random sampling. In particular, the roles of death and (random) loss may be interchangeable. By redefining the classification of events into deaths and losses, it may be possible to approach the same data from various points of view and thus to estimate the survivorship functions $P(t)$ that would be appropriate to various categories of events in the absence of the others. This is familiar enough; see for example [6], [13], [16].

1.2 *Nonparametric estimation.* Most general methods of estimation, such as maximum likelihood or minimum chi-square, may be interpreted as procedures for selecting from an admissible class of distributions one which, in a specified sense, best fits the observations. To estimate a characteristic (or parameter) of the true distribution one uses the value that the characteristic has for this best fitting distribution function. It seems reasonable to call an estimation procedure *nonparametric* when the class of admissible distributions from which the best-fitting one is to be chosen is the class of all distributions. (Wolfowitz [28] has used the term similarly in connection with the likelihood ratio in hypothesis testing). With a complete sample, it is easy to see that the sample distribution referred to in 1.1 is the nonparametric estimate on the maximum likelihood criterion. The same result is true of the product-limit estimate for incomplete samples, as will be demonstrated in Section 5.

The most frequently used methods of parametric estimation for distributions of lifetimes are perhaps the fitting of a normal distribution to the observations or their logarithms by calculating the mean and variance, and fitting an exponential distribution $e^{-t/\mu}dt/\mu$ by estimating the mean life μ alone. Such assumptions about the form of the distribution are naturally advantageous insofar as they are correct; the estimates are simple and relatively efficient, and a complete distribution is obtained even though the observations may be restricted in range. However, nonparametric estimates have the important functions of suggesting or confirming such assumptions, and of supplying the estimate itself in case suitable parametric assumptions are not known. An important property of these nonparametric estimates is that if the age scale is transformed from t to $t^* = f(t)$, where f is a strictly increasing function, then the corresponding estimated distribution functions are simply related by $\hat{F}^*(f(t)) = \hat{F}(t)$.

1.3 *Examples of the RS and PL estimates.* We will consider the following situation. A random sample of 100 items is put on test at the beginning of 1955; during the year 70 items die and 30 survive. At the end of the year, a larger sample is available and 1000 additional items are put on test. During 1956, 15 items from the first sample and 750 from the second die, leaving 15 and 250 survivors respectively. As of the end of 1956, it is desired to estimate the proportion $P(2)$ of items in the population surviving for two years or more.

The survival probabilities are supposed to depend on the age (the duration of the test) rather than on the calendar year, and hence the data are arranged as in Table 460.

TABLE 460

Samples	I	II
Initial numbers	100	1000
Deaths in first year of age	70	750
One-year survivors	30	250
Deaths in second year of age	15	
Two-year survivors	15	

This particular example is such that it is easy to form an estimate $P^*(2) = 15/100 = 0.15$ from the first sample alone. This is called the *reduced-sample* (RS) estimate because it ignores the 1000 items tested only during 1956. It is a legitimate estimate only when the reduced sample is itself a random sample; this will be the case only when (as assumed here) the observation limits (two years for the first sample, and one year for the second) are known for all items, deaths as well as losses. In the absence of this information, one would have no basis for discriminating among the 835 deaths observed before the age of two years. One cannot simply ignore the 250 losses at age one year; since only 15 items have survived for two years, $P(2)$ would then be estimated as $15/850 = .018$, an absurd result. The point is discussed further in 3.1 below.

We now inquire whether the second sample, under test for only one year, can throw any light on the estimate of $P(2)$. Clearly it will be necessary to assume that both samples have been drawn from the same population, an assumption that the RS estimate $P^*(2)$ avoided. At any rate, the estimates of $P(1)$ from the two samples, namely 0.30 and 0.25, are not sufficiently different to contradict the assumption. By combining the two samples, the estimate

$$\hat{P}(1) = P^*(1) = (30 + 250)/(100 + 1000) = .255$$

is obtained for $P(1)$. (In this case the RS has the same value as the other estimate to be discussed, the product-limit or PL.) This result exhausts the usefulness of the second sample for the present purposes; how does it help to estimate $P(2)$?

The answer is that there are advantages to using the first (the smaller) sample for estimating $P(2)/P(1)$, the conditional probability of survival for two years given survival for one year, rather than $P(2)$ itself. This estimate is

$$\hat{P}(2)/\hat{P}(1) = 15/30 = 0.50, \text{ whence}$$

$$\hat{P}(2) = 0.255 \times 0.50 = 0.127,$$

a very simple example of the *product-limit* (PL) estimate. The outstanding advantage of this strategy is that it works just as well if we are not privileged to know that the 750 deaths in the second sample had observation limits of one year, because these items are irrelevant to the estimation of $P(2)/P(1)$ in any

case. Other considerations for deciding between the two estimates will be set forth in Section 3.

The discussion of the PL estimate will be continued shortly, in Sections 2 and 3. Section 3 is equally concerned with the RS estimate, while Section 4 is devoted to the actuarial estimates. The remaining three sections consist principally of mathematical derivations. Though much older, the actuarial estimates are essentially approximations to the PL; they are products also, but typically they aim to reduce the number of factors by grouping. (Grouping may or may not be possible for the PL itself.) The distinguishing designation product-*limit* was adopted because this estimate is a limiting case of the actuarial estimates. It was proposed as early as 1912 by Böhmer [6] (referred to by Seal [26]), but seems to have been lost sight of by later writers and not further investigated.

1.4 *Notation.* The survival function

$$P(t) = \Pr(T > t), \tag{1c}$$

giving the population probability of surviving beyond t, will be used in place of the distribution function $F(t) = 1 - P(t)$ because of its convenience where the product-limit estimate and its actuarial approximations are concerned. In addition the following functions are defined:

$\hat{P}(t)$ = product-limit (PL) estimate of $P(t)$.
$P^*(t)$ = reduced-sample (RS) estimate of $P(t)$.
$n(t)$ = the number of items observed and surviving at age t, when deaths (but not losses) at t itself are subtracted off.
$N(t)$ = the expectation of $n(t)$, for fixed observation limits.
$N^0(t)$ = the number of items having observation limits L such that $L \geq t$. In practice this function is not necessarily known.

For the first reading of the paper it may be desirable to suppose that the death of one item and the loss of the same or any other item never occur at the same age, and never coincide with an age t at which any of the above functions are to be evaluated. This condition can always be met by fudging the ages a little when necessary. On the other hand, a regular user of the techniques will probably come to regard overt fudging as naive; he will prefer to formalize his notation and record-keeping by adopting the conventions already insinuated into the definitions of death and loss, $P(t)$, and $n(t)$ above. These conventions may be paraphrased by saying that deaths recorded as of an age t are treated as if they occurred slightly before t, and losses recorded as of an age t are treated as occurring slightly after t. In this way the fudging is kept conceptual, systematic, and automatic.

The convention that deaths precede losses in case of ambiguity is based on the following sequence of operations, which is clearly more efficient than the reverse sequence: Examine a group of items of age t_0, observe the number δ of deaths since the last examination, and then remove (or lose contact with) a number λ of the survivors. It may then be convenient simply to record t_0 as the age of death or loss of the $\delta + \lambda$ items, especially if t_0 is always an integral multiple of a fundamental time interval; in fact, however, the deaths will have preceded the losses.

The chief exception to the immediate applicability of this convention occurs when the losses are random but cannot affect items that have already died. Then the possible sequences of occurrence of deaths and losses between examination times (assumed to be close together) are approximately equally likely, and a reasonable compromise is to assume that half the losses in the interval precede and half follow the deaths, as in (4b) below. If the loss of an item is compatible with the possibility of its having died (unknown to the experimenter) between the time when it was last examined and the time of loss, an item lost in this way is effectively lost just after it was last examined, and the convention is entirely appropriate. The disappearance of individuals subject to medical follow-up is a case in point.

The remainder of the conventions concern the treatment of discontinuities in the functions listed above. Should the value assigned to $n(t)$, say, for an argument t for which it is discontinuous be the right-hand limit $n(t+0)=\lim n(t+h)$ as $h\to 0$ with $h>0$, the left-hand limit $n(t-0)$ (the same but with $h<0$), or something else? The superior expressiveness of the notation adopted is illustrated by the relations

$$\begin{aligned} n(t-0)-n(t) &= \text{the number of deaths at } t, \\ n(t)-n(t+0) &= \text{the number of losses at } t, \end{aligned} \tag{1d}$$

or equivalently by the formula (3a): $P^*(t)=n(t)/N^0(t)$, which otherwise would not be valid at discontinuities. Analogous conventions are adopted for all the above-mentioned functions of t, so that $P(t)$ and $\hat{P}(t)$ are right-continuous, $N^0(t)$ is left-continuous, and $n(t)$ and $P^*(t)$ are neither. Other advantages of the convention for $P(t)$ and its estimates are the following: (a) According to the sequence of operations assumed above, one may record deaths as of age t although they actually occurred slightly earlier. (b) It makes $P(0)=1$ if and only if no item dies at birth (age zero). This is convenient and natural.

One other possibility may be mentioned briefly. The assumption that $P(\infty)=0$, so that the lifetimes are finite with probability one, is necessary only for parts of Section 7 and for the calculation of a finite mean lifetime. However, in practice there is no apparent need to contradict the assumption either. If half of a sample dies in one day and the other half is still alive after 1000 days, one should still report $\hat{P}(1000)$ (not $\hat{P}(\infty)$) $=0.50$, since the argument 1000 is not an arbitrary large number, but the actual duration of the test.

2. THE PRODUCT-LIMIT ESTIMATE

2.1 *Definition and calculation.* Both the PL and the actuarial estimates of Section 4 are based on the following general procedure:

(a) The age scale is divided into suitably chosen intervals, $(0, u_1)$, (u_1, u_2), $\cdots$, as described below. (In the example of 1.3, there were only two such intervals, namely (0, 1) and (1, 2).)

(b) For each interval (u_{j-1}, u_j), one estimates $p_j=P_j/P_{j-1}$, the proportion of items alive just after u_{j-1} that survive beyond u_j.

(c) If t is a division point (it may be introduced specially if necessary), the proportion $P(t)$ in the population surviving beyond t is estimated by the product of the estimated p_j for all intervals prior to t.

If step (a) is left relatively arbitrary and approximations or parametric assumptions are accepted in step (b), one arrives at the actuarial estimates. The PL is obtained by selecting the intervals in (a) so that the estimation in (b) is a simple binomial, without any recourse to assumptions of functional form. The condition for this is that within each interval, deaths and losses be segregated in a known fashion. As a beginning, it may be assumed that no interval contains both deaths and losses. Then if the number under observation just after u_{j-1} is denoted by n_j, and δ_j deaths are observed in the interval (u_{j-1}, u_j), the estimate is clearly

$$\hat{p}_j = (n_j - \delta_j)/n_j = n_j'/n_j, \tag{2a}$$

where n_j' is the number under observation just after the δ_j deaths. However, if the interval contains only losses (but at least one item survives throughout the interval), the estimate is $\hat{p}_j = 1$.

In the product of conditional probabilities formed in step (c), unit factors may as well be suppressed; and we need not be concerned with the manner in which the losses are distributed among the intervals, so long as n_j and n_j' are correctly evaluated in (2a), and no losses occur at ages intermediate to the δ_j deaths. The situation is illustrated by the following scheme:

TABLE 463

No. of items	N		n_1		n_1'		n_2		n_2'		. . . .
No. of deaths or losses		λ_0		δ_1		λ_1		δ_2		λ_2	
Division points	$u_0 = 0$					u_1				u_2	

Here N is the initial number of items, and the braces join the numbers whose ratios are the conditional probabilities (2a). The numbers in the second line are the differences of those in the first, the λ's counting losses and the δ's deaths; some of these could be zero. The division points u_j are placed in the third line to show that the δ_j deaths occur between u_{j-1} and u_j, while u_j is located anywhere among the λ_j losses. The relation $n_j' \leq n(u_j) \leq n_{j+1}$ holds.

The PL estimate is now given by

$$\hat{P}(t) = \prod_{j=1}^{k} (n_j'/n_j), \quad \text{with} \quad u_k = t,\ n_j' = n_j - \delta_j. \tag{2b}$$

If the greatest observed lifetime t^* corresponds to a loss, (2b) should not be used with $t > t^*$; in this case $\hat{P}(t)$ can be regarded as lying between 0 and $\hat{P}(t^*)$, but is not more closely defined.

If it is desired to permit the entrance of items into the sample after the commencement of their lifetimes, this can be done by treating such entrances as "losses" that are counted negatively in λ_j. The same items can of course disappear again at a later age and so yield ordinary losses as well. It is assumed that nothing is known of the existence of any such item that dies before it becomes available for observation; that is, the observation is censored on the right but truncated on the left, in the terminology of Hald [17].

The form (2b) was selected for the PL estimate to give the minimum number of elementary factors and the maximum grouping of the observations. Nevertheless, the number of deaths δ_j in an interval can easily be as small as unity. The resulting estimate $\hat{p}_j$, though of limited value by itself, is none the less acceptable as a component of $\hat{P}(t)$. In fact, one is at liberty to take the intervals as short and as numerous as one pleases, and to regard each death as occupying an interval by itself. To specify the resulting expression, one relabels the N ages t_i of death or loss in order of increasing magnitude, and denotes them by $t_1' \leq t_2' \leq \cdots \leq t_N'$. Then

$$\hat{P}(t) = \prod_r [(N - r)/(N - r+1)], \tag{2c}$$

where r runs through those positive integers for which $t_r' \leq t$ and t_r' is the age of death (not loss). The cancellation of like integers in numerator and denominator where they occur reduces (2c) to (2b). If there are no losses, everything cancels except the first denominator N and the last numerator $n(t)$, say, and the PL reduces to the usual binomial estimate $n(t)/N$. (2c) shows that $\hat{P}(t)$ is a step-function which changes its value only at the observed ages of death, where it is discontinuous.

In analyzing data on lifetimes by the multiplication of conditional probabilities, one of the following three procedures will usually suffice:

(1) If the number of deaths is relatively small, these deaths may be arranged in order of age without grouping, and the numbers of losses in the intervening age intervals counted. The PL estimate is calculated by (2c).

(2) If (1) is too time-consuming but the number of distinct ages of loss is relatively small, these ages may be arranged in order, additional division points inserted as desired, and the numbers of deaths in the resulting age intervals counted. If some of these intervals are shorter than necessary and are found to contain no deaths, they can be combined with adjacent intervals. The PL estimate is calculated by (2b).

(3) If neither (1) nor (2) is compact enough, then division points are chosen without close consideration of the sample, deaths and losses are counted in each interval, and an actuarial approximation to the PL, such as (4b), is used.

As a miniature example of case (1), suppose that out of a sample of 8 items the following are observed:

Deaths at 0.8, 3.1, 5.4, 9.2 months.
Losses at 1.0, 2.7, 7.0, 12.1 months.

The construction of the function $\hat{P}(t)$ then proceeds as follows:

TABLE 464

u_j	n_j	n_j'	λ_j	$\hat{P}(u_j)$
0.8	8	7	2	7/8
3.1	5	4	0	7/10
5.4	4	3	1	21/40
9.2	2	1	0	21/80
(12.1)	1	1	1	21/80

Each value of $\hat{P}(u_j)$ is obtained by multiplying n_j'/n_j by the preceding value $\hat{P}(u_{j-1})$. The age 12.1 is recorded in the last line to show the point at which $\hat{P}(t)$ becomes undefined; since it is a loss time, the 12.1 is enclosed in parentheses. It is to be inferred from the table that $\hat{P}(5.3)=7/10$, for example. The third and fourth columns could be omitted since $n_j'=n_j-1$ (except in the last line, which corresponds to a loss) and $\lambda_j=n_j'-n_{j+1}$.

A rudimentary illustration of case (2) has already been given in 1.3. A little more elaborate example of a similar sort with $N=100$ is given in Table 465. Here 1.7, 3.6, and 5.0 are assumed to be the only ages at which losses occur; they are prescribed as division points. The other division points (1, 2, 3, 4) are selected at pleasure, with the object of interpolating additional points on the curve of $\hat{P}(t)$ vs. t. $\hat{N}_{Ej}$ is the effective sample size defined in (2j) below. In practice four columns headed u_j, n_j, n_j', $\hat{P}_j$ will suffice. From the table one infers that $.74<\hat{P}(2.5)<.87$, for example.

TABLE 465

Interval	Factor						
u_{j-1}, u_j	j	n_j	δ_j	λ_j	$\hat{p}_j$	$\hat{P}(u_j)$	$\hat{N}_{Ej}$
0 –1	1	100	3	0	97/100	.97	100
1 –1.7	2	97	5	20	92/97	.92	100
1.7–2	3	72	4	0	68/72	.87	88
2 –3	4	68	10	0	58/68	.74	83
3 –3.6	5	58	9	12	49/58	.63	80
3.6–4	6	37	6	0	31/37	.52	73
4 –5.0	7	31	15	16	16/31	.27	51

2.2 *Mean and variance of* $\hat{P}(t)$. The important facts here, derived in Section 6.1, are that $\hat{P}(t)$ is consistent and of negligible bias (unless excessive averaging is done; see Section 3.1), and that an asymptotic expression for its variance can be obtained. Like the estimate itself, the sample approximation to its variance proves to be independent of the limits of observation of items not actually lost. However, the variance derived from population values does depend on all the limits of observation, which are assumed to be fixed during the sampling.

It has been noted that if the greatest observed lifetime t^* corresponds to a loss, then for $t>t^*$, $\hat{P}(t)$ is undefined though bounded by 0 and $\hat{P}(t^*)$. Unless the probability of this ambiguous situation is quite small, however, a nonparametric estimate of $P(t)$ will not be very informative in any case. The ambiguity cannot occur unless the $N^0(t)$ items observable to t all die at ages less than t. The probability of this event is

$$[1-P(t)]^{N^0(t)} \leq e^{-N^0(t)P(t)} = e^{-N(t)} \tag{2d}$$

This is already less than 0.01 when $N(t)$ is only five.

It is shown in Section 6.1 that if one can supplement the ambiguous case by ascertaining the age of death of the item lost at t^*, or of one or more other randomly selected items alive at t^*, and defines $\hat{P}(t)$ for $t>t^*$ as $\hat{P}(t^*)$ times the survival function for the supplementary sample, then the expected value of

$\hat{P}(t)$ is precisely the population value $P(t)$. In practice this supplementation would often be neither feasible nor worthwhile, but with otherwise adequate data the resulting bias in one or a few samples will be too small to have any practical importance.

It will be shown in Section 6.1 that the variance of $\hat{P}(t)$ is given approximately by

$$V[\hat{P}(t)] \doteq P^2(t) \sum_1^k (q_j/N_j p_j), \tag{2e}$$

where the distinct limits of observation L_j' are now used as the division points; L'_{k-1} is the greatest preceding t; and $p_j = 1 - q_j = P(L_j')/P(L'_{j-1})$, with $L_0' = 0$, $L_k' = t$. After dividing (2e) by $P^2(t)$, one sees that the square of the coefficient of variation (CV) of $\hat{P}(t)$ is set equal to the sum of the squares of the CV's of the estimates of the p_j, the usual approximation for the variance of a product.

If the sample estimates are inserted in (2e) one obtains

$$\hat{V}[\hat{P}(t)] \doteq \hat{P}^2(t) \sum_1^k [\delta_j/n_j(n_j - \delta_j)] = \hat{P}^2(t) \sum_1^k \left(\frac{1}{n_j'} - \frac{1}{n_j}\right). \tag{2f}$$

A very similar formula was derived by Greenwood [15] and later by Irwin [19] in connection with actuarial estimates. It is easily verified that (2f) remains valid when the number of intervals is reduced to those used in (2b), or expanded to one interval for each death as in (2c). In the latter case it may be written

$$\hat{V}[\hat{P}(t)] \doteq \hat{P}^2(t) \sum_r [(N - r)(N - r + 1)]^{-1}, \tag{2g}$$

where r runs through the positive integers for which $t_r' \leq t$ and t_r' corresponds to a death.

In terms of integrals (2e) can be written

$$\begin{aligned} V[\hat{P}(t)] &\doteq P^2(t) \int_0^t \frac{|\,dP(u)\,|}{N^0(u)P^2(u)} \\ &\doteq P^2(t) \int_0^t \frac{|\,dP(u)\,|}{N(u)P(u)} \end{aligned} \tag{2h}$$

In case of ambiguity the first form should be referred to and interpreted in accordance with 1.4. If $P(u)$ alone is discontinuous within the range of integration, it should be regarded as a (continuous) independent variable, so that $\int P(u)^{-2} |\,dP(u)\,| = 1/P(u)$.

Since $N^0(t) \leq N^0(u) \leq N$, it is clear that (2h) is greater than the complete sample variance $P(t)[1-P(t)]/N$ but less than the reduced-sample variance $P(t)[1-P(t)]/N^0(t)$ of (3b). If losses are random and the instantaneous rates of death and loss among survivors are in the ratio 1 to ρ at all ages, then one has $E[N^0(u)^{-1}] \doteq [NP^\rho(u)]^{-1}$ for large samples, and (2h) reduces to

$$V[\hat{P}(t)] \doteq [P^{1-\rho}(t) - P^2(t)]/(1 + \rho)N \tag{2i}$$

It is often instructive to estimate an effective sample size $\hat{N}_E(t)$, which in the absence of losses would give the same variance $\hat{V}[\hat{P}(t)]$. Evidently

$$\hat{N}_E(t) = \hat{P}(t)[1 - \hat{P}(t)]/\hat{V}[\hat{P}(t)]. \tag{2j}$$

It can be shown that $\hat{N}_E(t)$ is nonincreasing as t increases, and that

$$n(t)/\hat{P}(t) \leq \hat{N}_E(t) \leq D(t)/[1 - \hat{P}(t)], \tag{2k}$$

where $D(t)$ is the number of deaths observed at ages not exceeding t. The upper bound was proposed as an approximation to $\hat{N}_E(t)$ by Cornfield [8]. The lower bound corresponds to the reduced-sample variance.

In Table 464 of 2.1 we had $\hat{P}(6) = 21/40 = 0.525$. By (2f) the variance of $P(6)$ is estimated as

$$\hat{V}[\hat{P}(6)] = (0.525)^2\left(\frac{1}{8 \times 7} + \frac{2}{5 \times 3}\right) = 0.042.$$

By (2g) the same result appears in the form

$$(0.525)^2\left(\frac{1}{7 \times 8} + \frac{1}{4 \times 5} + \frac{1}{3 \times 4}\right).$$

The effective sample size is estimated as

$$\hat{N}_E(6) = (0.525)(0.475)/0.042 = 6.0.$$

The bounds for $\hat{N}_E(6)$ in (2k) are 5.7 and 6.3. Values of $\hat{N}_E$ are also indicated in Table 465.

2.3 *Mean lifetime.* The PL estimate $\hat{\mu}$ of the mean lifetime μ is defined as the mean of the PL estimate of the distribution. It is well-known (and easily proved by integrating by parts) that the mean of a nonnegative random variable is equal to the area under the corresponding survivorship function. Hence

$$\hat{\mu} = \int_0^\infty \hat{P}(t)dt.$$

Of course, if $\hat{P}(t)$ is not everywhere determined, $\hat{\mu}$ is undefined. In cases where the probability of an indeterminate result is small, $\hat{P}(t)$ is practically unbiased and the same is true of $\hat{\mu}$.

If we "complete" $\hat{P}(t)$ in the example of Table 464 by following the longest observed individual (with $t_r' = 12.1$) to death at $t = 14.3$ we have

$$\begin{aligned}\hat{\mu} &= (1.000)(0.8) + (0.875)(3.1 - 0.8) + (0.700)(5.4 - 3.1) \\ &\quad + (0.525)(9.2 - 5.4) + (0.2625)(14.3 - 9.2) \\ &= 0.800 + 2.012 + 1.610 + 1.995 + 1.339 = 7.76.\end{aligned}$$

If $\hat{P}(t)$ were "completed" by setting $\hat{P}(t) = 0$ for the indeterminate range, the last term in the above sum would be replaced by $(0.2625)(12.1 - 9.2) = 0.761$, and $\hat{\mu}$ would be estimated as 7.18.

Of course, if the probability of an indeterminate result is high, there is no satisfactory way to estimate μ. In such cases Irwin [19] has suggested that in

place of estimating the mean itself, one should estimate the "mean life limited to a time L," say $\mu_{[L]}$. This is the mean of min (T_i, L), with L chosen at the investigator's convenience. Naturally, one would choose L to make the probability of an indeterminate result quite small. If one chooses to use this procedure he should give an estimate of $P(L)$ along with $\hat{\mu}_{[L]}$. If we take $L=10$ in our example,

$$\hat{\mu}_{[10]} = 0.800 + 2.012 + 1.610 + 1.995 + (0.2625)(10 - 9.2) = 6.63,$$

and $\hat{P}(10)=0.2625$.

In Section 6.2 an approximate formula is given for the variance of $\hat{\mu}$:

$$V(\hat{\mu}) \doteq \int_0^\infty \frac{A^2(t) \mid dP(t) \mid}{N(t)P(t)} = \int_0^\infty \frac{A^2(t) \mid dP(t) \mid}{N^\circ(t)P^2(t)}, \tag{2l}$$

where $A(t)=\int_t^\infty P(u)du$. Upon making the obvious substitutions we find, after some reduction, the following estimate of $V(\hat{\mu})$:

$$\hat{V}(\hat{\mu}) = \sum_r \frac{A_r^2}{(N-r)(N-r+1)} \tag{2m}$$

where r runs over those integers for which t_r corresponds to a death, and $A_r=\int_{t_r}^\infty \hat{P}(u)du$. If there are no losses,

$$A_r = \sum_{i=r+1}^{N} (t_i - t_r)/N,$$

and it can be shown that $\hat{V}(\hat{\mu})$ reduces to $\sum (t_i-\bar{t})^2/N^2$. This fact, plus the impossibility of estimating the variance on the basis of only one observed death, suggests that $(2m)$ might be improved by multiplication by $D/(D-1)$, where D is the number of deaths observed.

For our first estimate of $\hat{\mu}$ above we have

$$\begin{aligned} A_1 &= 2.012 + 1.610 + 1.995 + 1.339 = 6.956, \\ A_4 &= \quad 1.610 + 1.995 + 1.339 = 4.944, \\ A_5 &= \quad\quad 1.995 + 1.339 = 3.334, \\ A_7 &= \quad\quad\quad 1.339 = 1.339. \end{aligned}$$

(Obviously the A_r and $\hat{\mu}$ are best calculated in reverse order.) We then have for the estimated variance,

$$\hat{V}(\hat{\mu}) = \frac{(6.956)^2}{7\times 8} + \frac{(4.944)^2}{4\times 5} + \frac{(3.334)^2}{3\times 4} + \frac{(1.339)^2}{1\times 2} = 3.91$$

The estimated standard deviation of $\hat{\mu}$ is $\sqrt{3.91}=1.98$. If the factor $D/(D-1)=4/3$ is included, these results become 5.21 and 2.28 respectively.

If one is limited to grouped data as in Table 465 it is necessary to use actuarial-type assumptions (e.g., the trapezoidal rule) to estimate $\hat{\mu}$ and its variance. Thus, we may estimate the mean life limited to 5 time units as follows:

$$\begin{aligned} \hat{\mu}_{[5]} = (1/2)\,[&(1.00 + 0.97)(1.0 - 0.0) + (0.97 + 0.92)(1.7 - 1.0) \\ &+ \cdots + (0.52 + 0.27)(5.0 - 4.0)\,] = 3.76. \end{aligned}$$

We observe in conjunction with this estimate that the estimated proportion surviving the limit is $\hat{P}(5) = 0.27$.

3. THE REDUCED-SAMPLE ESTIMATE VS. THE PL

3.1 *Alternatives to the PL estimate.* The acceptable nonparametric estimation procedures known to the writers in the situation considered are variants of the PL and of the reduced-sample (RS) estimate defined by

$$P^*(t) = n(t)/N^0(t) = n(t)/[n(t) + D^0(t)]. \tag{3a}$$

Here $N^0(t) = n(t) + D^0(t)$ is the number of items with observation limits $\geq t$; of these, $D^0(t)$ die at ages $\leq t$ and $n(t)$ survive beyond t. A simple example of the RS estimate has been given in 1.3. Berkson and Gage [3] have called it the "ad-hoc" method.

If the lifetimes are independent of the observation limits, the reduced sample of $N^0(t)$ items will be a random sample and $P^*(t)$ a simple binomial estimate, unbiased, and with the variance

$$V[P^*(t)] = P(t)[1 - P(t)]/N^0(t). \tag{3b}$$

It has already been noted that this is equal to or larger than the approximation (2h) to the PL variance. As with the PL, the estimate

$$\mu^* = \int_0^\infty P^*(t)dt \tag{3c}$$

of the mean life may have to be truncated at the greatest of the observation limits. In Section 6.2 it is shown that the variance of μ^* is given exactly by

$$V(\mu^*) = 2\int_0^\infty \frac{1 - P(t)}{N^0(t)} \int_t^\infty P(u)dudt.$$

In practice one substitutes $P^*(t)$ for $P(t)$ in this formula and perhaps approximates the integrals.

Whereas the PL estimate has discontinuities only at observed ages of death, it is apparent from (3a) that the RS generally has discontinuities at losses and observation limits also. Furthermore, suppose that one of the observation limits L is such that all (or a sufficient number of) the corresponding items are observed to die prior to L, while other items survive and are observed beyond L. Then the decrease in $N^0(t)$ (without an equivalent decrease in $n(t)$) will cause $P^*(t)$ to *increase* as t increases through L. Thus $P^*(t)$, unlike $\hat{P}(t)$ and the true $P(t)$, is not necessarily monotonic decreasing. However, this appears to be a disadvantage only in a psychological sense. It does not seem advisable to avoid it, as one could, by a further reduction of the sample, basing the estimated $P(t)$ for all $t \leq$ a suitable L on the fixed sample of items having observation limits $\geq L$.

The expression $D^0(t) = N^0(t) - n(t)$ in (3a) is not the total number $D(t)$ of deaths observed prior to or at age t, but the (usually smaller) number of such deaths having observation limits $\geq t$. Since $P^*(t)$ is unbiased, the estimate obobtained by replacing $D^0(t)$ by $D(t)$ is too small on the average (except in

special cases where $D(t) = D^0(t)$ with probability one), as has been pointed out many times (e.g., [3] and [23]). Examples such as that of 1.3 show that the resulting bias can be very great, being in fact limited only by the necessity that the expectation be positive. Since $D^0(t)$ is often unknown while $D(t)$ is known, an investigator who is unacquainted with the PL or one of its approximations is all too likely to fall into the trap.

Another insidious but instructive characteristic of the illegitimate estimate based on $D(t)$ is the fact that it is simply the average of the PL estimates $\hat{P}(t)$ (those that are defined at t) obtained by regarding each item as a sample in its own right. For a sample of one item with observed lifetime t_1, $\hat{P}(t)$ is unity for $t < t_1$, and for $t \geq t_1$ is zero or undefined according as the observation ends with death or loss. This is also the RS estimate unless one insists that $P^*(t)$ is undefined for $t > L_1$ (the observation limit of the item), even though it may have been estimated as zero for smaller values of t.

The moral of this discussion is that arithmetic averages of independent PL estimates from very small samples will tend to be biased. On the other hand, the averaging of independent RS estimates is quite satisfactory, especially if the values of $N^0(t)$ are used as weights. To avoid bias one must at least assign zero weight when $N^0(t) = 0$, thus supplying each estimate with a limit of observation (the maximum of those attached to its items) beyond which it is regarded as undefined, even though a value of zero may be indicated at some earlier age.

To summarize, the advantages of the PL over the RS are the following:

(a) The observation limits need not be known for items observed to die.
(b) The sampling variance of the PL is usually a little smaller.
(c) The PL has fewer discontinuities, and its monotonicity may be comforting.

On the other hand, the RS has the following advantages over the PL:

(a) It is perfectly unbiased, and hence estimates from many small samples can be combined by (weighted) averaging.
(b) It facilitates the estimation of an isolated value of $P(t)$.
(c) It may be preferred in some cases of dependence between lifetimes and observation limits (see 3.2 below).

3.2 *Dependence of deaths and losses.* The assumption that the full lifetimes T_i are independent of the observation limits L_i is sometimes violated, as a result of a change in the population sampled, or the conditions leading to the event called death, or the method of sampling the population. For example, in a study of survival after an operation, a change in surgical technique five years before the data are analyzed will affect the survival times only of those with observation limit less than five years. When loss is due to unforeseen circumstances, such as patients moving out of the state, the possibilities for dependence are obvious. Merrell [23] and Sartwell [24] have emphasized this point.

The estimate of $P(t)$ from the sample as a whole, whether PL or RS, involves arbitrary assumptions whose danger is peculiar to the dependent case and hence easily overlooked. The PL estimate in effect assumes that for items having an

observation limit L less than t, the conditional survival probability $P(t)/P(L)$ is the same as that for items whose observation limits exceed t, while the RS estimate makes the same assumption concerning the absolute survival probability $P(t)$ itself. In many applications, the PL assumption may seem as plausible as any; on the other hand, the fact that the results are expressed in terms of absolute probabilities may lead one to prefer the RS assumption (when there is no feeling to the contrary), if only because its operation is more easily visualized. For example, the RS estimate $P^*(t)$ (apart from sampling fluctuations) can never fall outside the range of the true values for the populations sampled, because it is a weighted average, whereas in unfavorable cases the PL estimate $\hat{P}(t)$ may fall outside this range.

If prior observation limits are given for all the items and the sample is large enough, dependence may be inquired into by grouping the items according to these limits, making separate estimates of $P(t)$ for each group, and comparing the results. If some unanticipated losses occur in advance of the a priori limits, the PL method can be used within each group. The group estimates can then be averaged with the initial numbers of items as weights to give something similar to the RS estimate, if that is desired.

On the other hand, no dependence can be demonstrated if one's information is limited to the values of the $t_i = \min(L_i, T_i)$ and their classification as deaths or losses; in this case the observed rates of death and loss can always be represented by a model in which deaths and losses occur independently, as well as by many models in which this is not so. However, one usually has or can obtain other information that is more or less useful.

What one would like is to obtain for each item the value of an auxiliary variable V (which may be either quantitative or qualitative), such that the T_i and L_i are more nearly independent within subsamples defined in terms of V, than they are in the sample or population as a whole. This will generally be true if V is strongly related to the cause of loss, and independent of the cause of death, or vice versa. In this case the subsamples are formed so as to reduce the variability of the L_i (or the T_i) within them. Even if the auxiliary variable does not have the desired properties, it may be worthwhile to get an estimate of $P(t)$ by a different route for the sake of comparison. Also, the auxiliary variable may be of interest for its own sake, for example, classification of the items by starting date would be aimed at detecting a temporal change in the population. As pointed out in 3.1, the procedure in the general case is not satisfactory if the estimates depend on very few items.

The innocuous form of dependence that results from testing with replacement is considered briefly in Section 7.

4. ACTUARIAL ESTIMATES

4.1 *Estimates using only n, δ, λ.* It has already been indicated (in 2.1) that actuarial estimates, like the PL, are formed by multiplying together a sequence of estimates of conditional probabilities of survival through intervals $(0, u_1)$, (u_1, u_2), $\cdots$. Unlike the PL, the actuarial estimates will generally be somewhat dependent on the selection of these intervals. It remains to consider how a typical factor p of this sort can be estimated when the number of items n at the

beginning of the interval is known to be depleted by δ deaths and λ losses within the interval, but the order in which these occur is not known or used. The PL estimate is denoted by $\hat{p}$.

As in 2.1, the estimates

$$\bar{p} = \frac{n - \delta}{n}, \qquad \underline{p} = \frac{n - \lambda - \delta}{n - \lambda} \tag{4a}$$

would be used if all the deaths were known to precede all the losses, and if all losses preceded all deaths, respectively. Evidently $\underline{p} \leq \hat{p} < \bar{p}$. If $\hat{p}$ is not known, another intermediate value is clearly supplied by the well-known "adjusted-observed" estimate (cf. [3], [12], [18], [22])

$$p^{(1)} = \frac{n - \lambda/2 - \delta}{n - \lambda/2}. \tag{4b}$$

This can be recommended for its simplicity even though

$$\frac{n - \lambda/2 - \delta}{n - \lambda/2} \cdot \frac{n - \delta/2 - \lambda}{n - \delta/2} \neq \frac{n - \delta - \lambda}{n} \tag{4c}$$

Here the second fraction is the corresponding estimate of the probability of not being lost (if losses are also random events), and the third fraction is the observed probability of escaping both death and loss. The only other estimate in this section that fails to give equality here is $p^{(4)}$ below.

The "joint risk" estimate (cf. [6] and [12]) is

$$p^{(2)} = \left[\frac{n - \delta - \lambda}{n}\right]^{\delta/(\delta+\lambda)} \tag{4d}$$

It is the maximum likelihood estimate when the losses are random, the instantaneous event-rate for losses is a constant times that for deaths, and the only data given are the values of n, δ, and λ. The above estimates satisfy the relation $\underline{p} \leq p^{(2)} \leq p^{(1)} \leq \bar{p}$.

One can obtain the PL estimate $\hat{p}$ from any of the foregoing by dividing the given interval into smaller intervals no one of which contains both deaths and losses. Since there are reasons (full use of information, absence of arbitrary assumptions, essential uniqueness of the estimate) for preferring small intervals to large, it seems reasonable to regard $\hat{p}$ as the standard to which the large-interval actuarial methods are approximations.

The error incurred in an actuarial estimate of p can be attributed to two sources, namely, the sampling error of $\hat{p}$, reflecting the fact that δ is a random variable, and the discrepancy between $p^{(1)}$ and $\hat{p}$, reflecting the unknown arrangement of deaths and losses within the interval. These sources of error will be compared by calculating the variance of $\log \hat{p}$ in each case, on the assumption that n is large compared with δ and λ, and that factors of the form $1+0(n^{-2})$ can be neglected. On this basis the effect of variation in δ is indicated sufficiently by the square of the coefficient of variation of $p^{(1)}$, namely

$$\delta/(n - \lambda/2)(n - \delta - \lambda/2) \doteq 4\delta(2n - \delta - \lambda)^{-2}. \tag{4e}$$

To evaluate the second source of error we note first that the total range of possible variation in $\log \hat{p}$ (given δ and λ) is by (4a)

$$\log \bar{p} - \log \underline{p} \doteq 4\delta\lambda(2n - \delta - \lambda)^{-2}. \tag{4f}$$

However, if the arrangement is random, the extreme values may be unlikely. Therefore it will now be assumed that all permutations of the δ deaths and λ losses are equally likely (which happens to be a consequence of the proportional event-rates leading to (4d)). In this case $\log \hat{p}$ is minus the sum of a random sample of δ of the $\delta+\lambda$ numbers

$$\log\left(1+\frac{1}{n-1}\right), \log\left(1+\frac{1}{n-2}\right), \cdots, \log\left(1+\frac{1}{n-\delta-\lambda}\right).$$

Sampling theory and the Euler-Maclaurin theorem lead to the approximate variance

$$4\delta\lambda(\delta + \lambda + 1)/3(2n - \delta - \lambda)^4. \tag{4g}$$

Dividing the sampling variance (4e) into the square of half the range (4f) of the grouping error gives

$$\delta\lambda^2(2n - \delta - \lambda)^{-2} \tag{4h}$$

which may be either >1 or <1. Dividing (4e) into the grouping variance (4g) gives

$$\lambda(\delta + \lambda + 1)/3(2n - \delta - \lambda)^2, \tag{4i}$$

which is always <1 (when $n>0$).

The above results suggest that rather large intervals could be used if enough was known about the mechanisms of death and loss, and that the principal source of error is the probable failure of the various permutations of deaths and losses within the interval to be equally likely, or to fit any other scheme that may be assumed. Evidence on this point could be obtained by examining the permutations that actually occur. For example, this is what is done in using (2a). Otherwise one should probably require at the least that (4h) be suitably small. This still assumes that grouping errors, like the sampling errors, are independent from interval to interval. If not, one is thrown back on the ultimate in conservatism, which is to assert only that $\prod_j \underline{p}_j \leq \hat{P} \leq \prod_j \bar{p}_j$.

4.2 *Estimates using average ages.* Several estimates will now be considered that make use of the individual ages of death and loss, although arbitrary age intervals continue to be used. The individual ages enter via the average age Δ of the δ deaths, and the average age Λ of the λ losses, both measured from the beginning of the age interval under consideration, which is of length h.

The first estimate assumes nothing about the losses, but assumes that the instantaneous death-rate is constant throughout the interval. The total exposure to the risk of death in the interval is then $(n-\delta-\lambda)h+\delta\Delta+\lambda\Lambda$, the maximum-likelihood estimate of the death-rate is $\delta/[(n-\delta-\lambda)h+\delta\Delta+\lambda\Lambda]$, and the corresponding estimate of the survival probability p is

$$p^{(3)} = \exp\left(-\delta/[n - \delta - \lambda + (\delta\Delta + \lambda\Lambda)/h]\right). \tag{4j}$$

By using $p^{(3)}$, Harris, Meier, and Tukey [18] have been able to give an estimation procedure for the more general and more difficult case in which the age of death is never known exactly, but only known to fall in some interval, semi-infinite or finite, bounded by one or two observation points which vary from item to item.

The PL estimate $\hat{p}$ can be derived from $p^{(3)}$ as well as from $p^{(1)}$ or $p^{(2)}$. Although the limit of the product of the estimates $p^{(3)}$ for the subintervals is not quite unique, the average value of log $p^{(3)}$ corresponding to δ deaths occurring at the same age is

$$-\delta \int_0^1 \frac{du}{n - \delta(1 - u)} = \log \frac{n - \delta}{n}, \tag{4k}$$

in agreement with $\hat{p} = (n-\delta)/n$. Here $\Delta(=uh)$ has been regarded as uniformly distributed between 0 and h. However, in other respects the behavior of $p^{(3)}$ does not parallel that of $\hat{p}$. If the position of one or more of the deaths is shifted toward the end of the interval, $p^{(3)}$ is increased, while $\hat{p}$ is decreased (or else not changed). Again, $p^{(3)}$ does not necessarily agree with $\hat{p}$ in the simple case of no losses. Finally, the extreme values of $p^{(3)}$, obtained by putting $\Delta = \Lambda = 0$ or h, fall outside the interval $\underline{p}$, $\bar{p}$ within which $\hat{p}$ is confined; in fact one can have $p^{(3)} < \underline{p}$ while $\hat{p} = \bar{p}$, or $p^{(3)} > \bar{p}$ while $\hat{p} = \underline{p}$.

The following estimates have been constructed with the object of using the mean ages Δ and Λ to simulate the behavior of $\hat{p}$:

$$p^{(4)} = (n - c\lambda - \delta)/(n - c\lambda), \tag{4l}$$

$$p^{(5)} = \left[\frac{n - \lambda - \delta}{n - \lambda}\right]^c \left[\frac{n - \delta}{n}\right]^{1-c} \tag{4m}$$

where

$$\begin{aligned} c &= 1/2 + (\Delta - \Lambda)/h \quad \text{for} \quad |\Delta - \Lambda| \leq h/2, \\ &= 0 \qquad\qquad\quad \text{for} \quad \Delta - \Lambda \leq -h/2, \\ &= 1 \qquad\qquad\quad \text{for} \quad \Delta - \Lambda \geq h/2. \end{aligned} \tag{4n}$$

The value of c is immaterial when $\delta = 0$ or $\lambda = 0$.

These estimates are two different averages of the bounding estimates $\underline{p}$ and $\bar{p}$, taken with weights c and $1-c$ respectively. The quantity c is so defined that the appropriate one of the bounding estimates is assigned as soon as $|\Delta - \Lambda| \geq h/2$. This is intuitively reasonable, and is justified analytically by the fact that log $p^{(4)}$ and log $p^{(5)}$ have essentially the variance (4g) of log $\hat{p}$, if each death and each loss is uniformly and independently distributed in the interval. To estimate this variance, one can multiply

$$\text{var } c \doteq \text{var } [(\Delta - \Lambda)/h] = (\delta + \lambda)/12\delta\lambda$$

by the square of the range (4f), since (4f) gives the change in log $p^{(4)}$ and log $p^{(5)}$ corresponding to a unit change in c.

Experience may indicate which estimate is preferable in a given type of situation. If relevant experience is lacking, it would seem advisable to rely

primarily on adequately small intervals, and to minimize the computation time by using simple estimates. This suggests that $p^{(1)}$ be used when the deaths and losses are thought to be arranged at random within the interval, $\bar{p}$ when the interval is chosen so that most of the deaths precede most of the losses, and $\underline{p}$ when losses precede deaths.

This completes the material that seems to be most important for applications. The remaining three sections are devoted primarily to mathematical derivations.

5. MAXIMUM LIKELIHOOD DERIVATION OF THE *PL*

In accordance with 1.2 the method of maximum likelihood is viewed as a means for selecting the best fitting distribution from a class of admissible distributions. It will now be shown that if all possible distributions are admitted, the product-limit is the one for which the sample likelihood is a maximum.

Let $T_1<T_2<\cdots, <T_k$ denote the distinct ages of death, with δ_j deaths observed at T_j. Let λ_j denote the number of losses in the interval $[T_j, T_{j+1})$, including any losses at T_j but none at T_{j+1}. (Here $j=0, 1, \cdots, k$; $T_0=0$; $T_{k+1}=\infty$.) The ages of the λ_j losses are denoted by $L_i^{(j)}$, $i=1, 2, \cdots, \lambda_j$. Only in this section, $P(t)$ will represent not the "true" survivorship but rather an arbitrary one that is to be determined so as to maximize the likelihood. The likelihood may now be written

$$\begin{aligned}&\left[\prod_{i=0}^{\lambda_0} P(L_i^{(0)})\right]\\ &\cdot[P(T_1-0)-P(T_1)]^{\delta_1}\left[\prod_{i=1}^{\lambda_1} P(L_i^{(1)})\right]\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ &\cdot[P(T_k-0)-P(T_k)]^{\delta_k}\left[\prod_{i=1}^{\lambda_k} P(L_i^{(k)})\right].\end{aligned} \tag{5a}$$

Here $P(T_j-0)$ includes, whereas $P(T_j)$ excludes, the probability of death at age T_j exactly.

Clearly (5a) will be maximized by making the $P(L_i^{(j)})$ and the $P(T_j-0)$ as large, and the $P(T_j)$ as small, as is consistent with the monotonicity of the function and the assumption $T_j\leq L_i^{(j)}<T_{j+1}$. This means that $P(L_i^{(0)})=1$ and

$$P(T_j) = P(L_i^{(j)}) = P(T_{j+1}-0) \tag{5b}$$

for all i and j. If the common value is denoted by P_j, (5a) reduces to

$$\prod_{j=1}^{k}(P_{j-1}-P_j)^{\delta_j}P_j^{\lambda_j} \quad \text{with} \quad P_0=1. \tag{5c}$$

Now writing $p_j=1-q_j=P_j/P_{j-1}$ gives $P_j=p_1p_2\cdots, p_j$ and $P_{j-1}-P_j=p_1\cdots p_{j-1}q_j$. With these substitutions (5c) becomes

$$\prod_{j=1}^{k} p_j^{n_j-\delta_j}q_j^{\delta_j}, \tag{5d}$$

where

$$n_j = N - \sum_{i=0}^{i=j-1} \lambda_i - \sum_{i=1}^{i=j-1} \delta_i = \sum_{i=j}^{i=k} (\lambda_i + \delta_i), \quad \text{or} \quad n(T_j - 0)$$

in the notation of 1.4. In this form each factor with a fixed index j is maximized individually by the binomial estimate $p_j = \hat{p}_j = (n_j - \delta_j)/n_j$, in agreement with the PL in (2a). If $\lambda_k > 0$, so that the greatest observed lifetime t^* corresponds to a loss, then $\hat{P}(t^*) > 0$ and the likelihood of the sample is independent of the values of $P(t)$ for $t > t^*$. Thus $\hat{P}(t)$ and the maximum likelihood estimate are indeterminate under the same circumstances, and the identity between them is complete.

6. MEANS AND VARIANCES

In this section the mean of $\hat{P}(t)$ and the variances of $\hat{P}(t)$ and of the mean-life estimates $\hat{\mu}$ and μ^* are derived. The observation limits are regarded as fixed from sample to sample. In principle, unconditional results for random loss times could be obtained by integration.

6.1 *The PL estimate* $\hat{P}(t)$. Let $L_1 < L_2 < \cdots < L_{k-1}$ be the distinct observation limits that are less than the age $t = L_k$ at which $P(t)$ is being estimated. Let $n_{j+1} = n(L_j + 0)$ be the number of survivors[1] observed beyond L_j; $N_{j+1} = En_{j+1} = N^0(L_j + 0)P(L_j)$; $p_j = 1 - q_j = P(L_j)/P(L_{j-1})$; and δ_j the number of deaths observed in the interval $(L_{j-1}, L_j]$, excluding nonzero values of the left endpoint, with $L_0 = 0$. Then

$$\hat{P}(t) = \prod_{j=1}^{k} \hat{p}_j \quad \text{with} \quad \hat{p}_j = (n_j - \delta_j)/n_j. \tag{6a}$$

Let E_j denote a conditional expectation for $n_1, \cdots, n_j$ (and hence also $\delta_1, \cdots, \delta_{j-1}$) fixed. Then $E_j\hat{p}_j = p_j$, provided that $n_j > 0$. If the condition $n_j > 0$ could be ignored, one could write

$$\begin{aligned} E\hat{P}(t) &= E[\hat{p}_1 \cdots \hat{p}_{k-1}E_k\hat{p}_k] = E[\hat{p}_1 \cdots \hat{p}_{k-1}p_k] \\ &= p_k E[\hat{p}_1 \cdots \hat{p}_{k-2}E_{k-1}\hat{p}_{k-1}] \\ &= \cdots = p_k p_{k-1} \cdots p_1 = P(t), \end{aligned} \tag{6b}$$

so that $\hat{P}(t)$ would be an unbiased estimate. The flaw in this demonstration is the indeterminacy resulting when (for some j) $n_{j+1} = 0$ but $\hat{P}(L_j) > 0$, an event whose probability is bounded by (2d), and which could be obviated in principle by obtaining one or more supplementary observations. In the derivations of approximate formulas that follow, any bias that $\hat{P}(t)$ may have is neglected.

Both Greenwood [15] and Irwin [19] when giving formulas for the variance of an actuarial estimate of $P(t)$ actually treat a special case in which the actuarial estimator and $\hat{P}(t)$ coincide. Translated into our notation their common argument is the following.

Suppose we consider conditional variances for n_j held fixed at the values N_j.

[1] The expression $L_j + 0$ means that losses at L_j itself have been subtracted off.

Then the $\hat{p}_j$ may be treated as independent quantities and

$$E[\hat{P}^2(t)] = \prod_{j=1}^{k}\left(p_j^2 + \frac{p_j q_j}{N_j}\right) = p_1^2 \cdots p_k^2 \prod_{j=1}^{k}\left(1 + \frac{q_j}{N_j p_j}\right)$$
$$= P^2(t) \prod_{j=1}^{k}\left(1 + \frac{q_j}{N_j p_j}\right). \tag{6c}$$

If we ignore terms of order N_j^{-2} we have as an approximation for the variance

$$V[\hat{P}(t)] \doteq P^2(t) \sum_{j=1}^{k} \frac{q_j}{N_j p_j}. \tag{6d}$$

This last expression we will refer to as "Greenwood's formula." The calculation of the variance with $n_1, \cdots, n_k$ fixed can apparently be justified only on the ground that it doesn't matter a great deal; the fixed n_j are unrealistic, and inconsistent with the previous assumption of a sample of fixed size and fixed limits of observation.

The authors have therefore rederived (6d) by means of successive conditional expectations. The procedure is to verify by induction the approximate relation

$$E_{j+1}[\hat{P}^2(t)] \doteq \hat{p}_1^2 \cdots \hat{p}_j^2\left(1 + \sum_{i=j+1}^{i=k} \frac{q_i}{p_i E_{j+1} n_i}\right) p_{j+1}^2 \cdots p_k^2 \tag{6e}$$

which reduces to (6d) when $j=0$. Applying E_j to (6e) has the effect of replacing j by $j-1$ provided that one sets

$$E_j \frac{\hat{p}_j^2}{E_{j+1} n_i} \doteq E_j\left[\hat{p}_j \cdot \frac{E_j \hat{p}_j}{E_j E_{j+1} n_i}\right]$$
$$\doteq p_j^2 / E_j n_i. \tag{6f}$$

wherein $\hat{p}_j / E_{j+1} n_i$ has been replaced by the ratio of the expectations of numerator and denominator. The fact that $\hat{p}_j$ and $E_{j+1} n_i$ are positively correlated improves the approximation.

In the special case in which $k=2$ (all observation limits less than t are equal to L_1) and any indeterminacy is resolved by one supplementary observation, the exact variance of $\hat{P}(t)$ can be calculated to be

$$V[\hat{P}(t)] = P^2(t)\left(\frac{q_1}{N p_1} + \frac{q_2}{N' p_1 p_2}\left[1 + \left(1 - \frac{N'}{N}\right)^2 \frac{N' p_1 - [N', p_1]}{[N', p_1]}\right.\right.$$
$$\left.\left. + \frac{N'}{N}\left(1 - \frac{N'}{N}\right)\left(\frac{q_1}{[N', p_1]} - 2q_1^{N'}\right)\right]\right), \tag{6g}$$

where N is the total sample size; $N' = N^0(t)$, the number of items observable to t or beyond; and

$$[N', p_1] = 1/E(X^{-1}) \tag{6h}$$

with $X = \max(Y, 1)$ and Y distributed binomially (N', p_1). The methods used by Stephan [27] show that $[N', p_1] = N' p_1 + 0(1)$ for N' large. Thus the term

in square brackets is of order $1+0(1/N')$, and (6g) approaches Greenwood's formula (6d) for large N'.

6.2 *Covariances and mean lifetimes.* The method of Irwin [19] can be used to derive the variance of $\hat{\mu} = \int_0^\infty \hat{P}(t)dt$ by writing

$$\begin{aligned} E(\hat{\mu}^2) &= E\int_0^\infty \int_0^\infty \hat{P}(u)\hat{P}(v)dvdu \\ &= 2\int_0^\infty \int_u^\infty E[\hat{P}(u)\hat{P}(v)]dvdu. \end{aligned} \tag{6i}$$

If $u<v$ and L_h and L_k denote u and v respectively, then apart from bias due to indeterminate cases one has

$$\begin{aligned} E[\hat{P}(u)\hat{P}(v)] &= E[\hat{P}^2(u)E_{h+1}(\hat{p}_{h+1}\hat{p}_{h+2}\cdots\hat{p}_k)] \\ &\doteq p_{h+1}p_{h+2}\cdots p_k E[\hat{P}^2(u)] \\ &\doteq [P(v)/P(u)]E[\hat{P}^2(u)] \doteq P(u)P(v)[1+U(u)], \end{aligned} \tag{6j}$$

where by (2h) and (2g)

$$U(u) = \int_0^u \frac{|dP(t)|}{N(t)P(t)} \doteq \sum [(N-r)(N-r+1)]^{-1};$$

the summation is over deaths at ages not exceeding u, and gives the sample estimate. Thus one has approximately

$$\begin{aligned} V(\hat{\mu}) &\doteq 2\int_0^\infty \int_u^\infty P(u)P(v)U(u)dvdu \\ &= 2\int_0^\infty A(u)P(u)U(u)du = \int_0^\infty A^2(u)dU(u), \end{aligned} \tag{6k}$$

where $A(u) = \int_u^\infty P(v)dv$. The result is discussed in 2.3.

To obtain analogous (but in this case exact) results for the RS estimates $P^*(t)$ and μ^*, let $u<v$ as before, and split off two independent subsamples consisting of the $N' = N^0(v)$ items observable to v or beyond, and the $N'' = N^0(u) - N^0(v)$ additional items observable to u or beyond, but not to v. Let the first subsample yield δ' deaths in $(0, u)$ and ϵ' in (u, v), while the second yields δ'' deaths in $(0, u)$. Also let $P_1 = P(u)$ and $P_2 = P(v)$. Then

$$\begin{aligned} &E[P^*(u)P^*(v)] \qquad (\text{with } u \le v) \\ &= E[(N'+N''-\delta'-\delta'')(N'-\delta'-\epsilon')]/N'(N'+N'') \\ &= E[(N'+N''-\delta'-\delta'')(N'-\delta')]P_2/P_1N'(N'+N'') \\ &= [E(N'-\delta')^2 + E(N''-\delta'')\cdot E(N'-\delta')]P_2/P_1N'(N'+N'') \\ &= [N'^2P_1^2 + N'P_1(1-P_1) + N'N''P_1^2]P_2/P_1N'(N'+N'') \\ &= P(u)P(v) + P(v)[1-P(u)]/N^0(u). \end{aligned} \tag{6l}$$

Substituting this result in the analogue of (6i) and subtracting μ^2 gives $V(\mu^*)$ as in (3d).

7. CONSISTENCY; TESTING WITH REPLACEMENT

The RS estimate $P^*(t)$, being binomial, is of course consistent. In view of the approximate nature of the formula for the variance of $\hat{P}(t)$, however, a proof of its consistency is in order. In fact, by giving closer attention to the errors of the approximations in the variance derivation, it can be shown that $\hat{P}(t)$ does have the limit $P(t)$ in probability provided only that $N^0(t)\to\infty$. However, the proof is too lengthy to be given here. Of course, the consistency is obvious enough in the special case in which the number k of conditional probabilities to be estimated remains bounded as $N^0(t)\to\infty$.

Thus far the independence of deaths and losses has been assumed. The question of consistency may also be raised in the differing context of "testing with replacement." (Cf. [7], [10], [11], [14].) This is the common life-testing situation in which a fixed number of items, say ν of them, are always under test; when one dies it is replaced by another. The only preassigned constants are the duration of the test, S, and the number, ν, of items on test at all times. The number of losses is then fixed at ν, one for each item on test at the conclusion of the experiment, while the number of deaths observed, and hence the total number of items, is a random variable. The first ν items on test have S as their limit of observation; if one of these dies at age t, with $t<S$, its successor has $S-t$ as its limit of observation, and so on. Consistency will mean an approach to $P(t)$ as $\nu S\to\infty$, which implies that either $\nu\to\infty$ or $S\to\infty$ (or both).

The RS estimate will be considered first. Let $T_1, T_2, \cdots$, be an infinite sequence of independent lifetimes with survivorship $P(t)$ and let $P_k(t)$ be the proportion of $T_1, \cdots, T_k$ that exceed t. By the strong law of large numbers, $\lim P_k(t)=P(t)$ with probability one as $k\to\infty$, for any specified t. Now the RS estimate $P^*(t)$ (which exists only for values of t in the interval 0, S) can be represented as a value of $P_k(t)$, with k a random variable. This can be done by assigning the lifetimes $T_1, \cdots, T_\nu$ to the ν items initially tested. If some of these items have replacements whose limits of observation $S-T_i$ $(i=1, \cdots, \nu)$ are not less than t, then succeeding lifetimes $T_{\nu+1}, T_{\nu+2}, \cdots$, are assigned to these replacements. The process is repeated until no more observation intervals of at least the duration t are left. To prove $P^*(t)$ consistent, it only remains to prove that for any M, $\Pr(k>M)\to 1$ as $\nu S\to\infty$. If $\nu\to\infty$ the result is evident. If $S\to\infty$, then even for $\nu=1$,

$$\begin{aligned}\Pr(k>M) &= \Pr(T_1+\cdots+T_M\le S-t)\\ &\ge \Pr[T_i\le (S-t)/M \text{ for } i=1,\cdots,M] \qquad (7a)\\ &= \{1-P[(S-t)/M]\}^M.\end{aligned}$$

As $S\to\infty$ the last expression has the limit unity and consistency is proved, provided that $P(\infty)=0$ (eventual death is certain).

When $S\to\infty$, ν is fixed, and $P(\infty)=0$, the consistency of the PL estimate $\hat{P}(t)$ follows from the relation (cf. (4a))

$$\frac{n(t)}{N^0(t)+D'}\le \hat{P}(t)\le \frac{n(t)+\nu}{N^0(t)+D'+\nu}=\frac{n(t)+\nu}{N} \qquad (7b)$$

since the number of losses is ν, $P^*(t) = n(t)/N^0(t)$ is consistent, $N^0(t) \to \infty$, and the number of deaths D' having observation limits less than t is a random variable that does not depend significantly on S. $\hat{P}(t)$ is undoubtedly consistent also when S is fixed and $\nu \to \infty$, but a rigorous proof of this has not been constructed.

ACKNOWLEDGMENT

The viewpoint adopted in this paper owes much to discussions with John W. Tukey. It also incorporates a number of suggestions kindly forwarded by readers of a preliminary draft.

REFERENCES

[1] Bailey, W. G., and Haycocks, H. W., "A synthesis of methods of deriving measures of decrement from observed data," *Journal of the Institute of Actuaries*, 73 (1947), 179–212.

[2] Bartlett, M. S., "On the statistical estimation of mean lifetime," *Philosophical Magazine* (Series 7), 44 (1953), 249–62.

[3] Berkson, J., and Gage, R. P., "Calculation of survival rates for cancer," *Proceedings of the Staff Meetings of the Mayo Clinic*, 25 (1950), 270–86.

[4] Berkson, J., and Gage, R. P., "Survival curve for cancer patients following treatment," *Journal of the American Statistical Association*, 47 (1952), 501–15.

[5] Berkson, J., "Estimation of the interval rate in actuarial calculations: a critique of the person-years concept," (Summary) *Journal of the American Statistical Association* 49 (1954), 363.

[6] Böhmer, P. E., "Theorie der unabhängigen Wahrscheinlichkeiten," Rapports, Mémoires et Procès-verbaux de Septième Congrès International d'Actuaires, Amsterdam, 2 (1912), 327–43.

[7] Brown, G. W. and Flood, M. M., "Tumbler mortality," *Journal of the American Statistical Association*, 42 (1947), 562–74.

[8] Cornfield, J., "Cancer illness among residents in Atlanta, Georgia," Public Health Service Publication No. 13, Cancer Morbidity Series No. 1, 1950.

[9] Davis, D. J., "An analysis of some failure data," *Journal of the American Statistical Association* 47 (1952), 113–150.

[10] Epstein, Benjamin, and Sobel, Milton, "Lifetesting," *Journal of the American Statistical Association*, 48 (1953), 486–502.

[11] Epstein, Benjamin, and Sobel, Milton, "Some theorems relevant to lifetesting from an exponential distribution," *Annals of Mathematical Statistics*, 25 (1954), 373–81.

[12] Fix, Evelyn, "Practical implications of certain stochastic models on different methods of follow-up studies," paper presented at 1951 annual meeting of the Western Branch, American Public Health Association, November 1, 1951.

[13] Fix, Evelyn, and Neyman, J., "A simple stochastic model of recovery, relapse, death and loss of patients," *Human Biology*, 23 (1951), 205–41.

[14] Goodman, L. A., "Methods of measuring useful life of equipment under operational conditions," *Journal of the American Statistical Association*, 48 (1953) 503–30.

[15] Greenwood, Major, "The natural duration of cancer," Reports on Public Health and Medical Subjects, No. 33 (1926), His Majesty's Stationery Office.

[16] Greville, T. N. E., "Mortality tables analyzed by cause of death," *The Record*, American Institute of Actuaries, 37 (1948), 283–94.

[17] Hald, A., "Maximum likelihood estimation of the parameters of a normal distribution which is truncated at a known point," *Skandinavisk Aktuarietidskrift*, 32 (1949) 119–34.

[18] Harris, T. E., Meier, P., and Tukey, J. W., "Timing of the distribution of events between observations," *Human Biology*, 22 (1950), 249–70.

[19] Irwin, J. O., "The standard error of an estimate of expectational life," *Journal of Hygiene*, 47 (1949), 188–9.

[20] Jablon, Seymour, "Testing of survival rates as computed from life tables," unpublished memorandum, June 14, 1951.
[21] Kahn, H. A., and Mantel, Nathan, "Variance of estimated proportions withdrawing in single decrement follow-up studies when cases are lost from observation," unpublished manuscript.
[22] Littell, A. S., "Estimation of the T-year survival rate from follow-up studies over a limited period of time," *Human Biology*, 24 (1952), 87–116.
[23] Merrell, Margaret, "Time-specific life tables contrasted with observed survivorship," *Biometrics*, 3 (1947), 129–36.
[24] Sartwell, P. E., and Merrell, M., "Influence of the dynamic character of chronic disease on the interpretation of morbidity rates," *American Journal of Public Health*, 42 (1952), 579–84.
[25] Savage, I. R., "Bibliography of nonparametric statistics and related topics," *Journal of the American Statistical Association*, 48 (1953), 844–906.
[26] Seal, H. L., "The estimation of mortality and other decremental probabilities," *Skandinavisk Aktuarietidskrift*, 37 (1954), 137–62.
[27] Stephan, F. F., "The expected value and variance of the reciprocal and other negative powers of a positive Bernoullian variate," *Annals of Mathematical Statistics*, 16 (1945), 50–61.
[28] Wolfowitz, J., "Additive partition functions and a class of statistical hypotheses," *Annals of Mathematical Statistics*, 13 (1942), 247–79.

INADMISSIBLE SAMPLES AND CONFIDENCE LIMITS*

HOWARD L. JONES
Illinois Bell Telephone Company

A simple procedure is proposed for dealing with situations where a possibility exists that a sample to be randomly selected may turn out to be unsatisfactory in some respect. The procedure consists essentially in computing confidence limits as if the specified risk were $\alpha_A(1-p_N)$, where α_A is the risk we are actually willing to run, and p_N is the a priori probability that a selected sample will not be considered to be satisfactory.

INTRODUCTION

CONFIDENCE limits computed from sample observations in the usual manner are perfectly valid, of course, if the underlying assumptions are correct. One of these assumptions is that, when the observations are taken, there is no possibility that any or all of them will be discarded as unsatisfactory. On rare occasions, however, probability samples are selected that are so unusual in some respect that a question arises as to whether the sample should be retained or whether a new sample should be selected.

In a previous paper [3], an actual situation was mentioned (p. 64) where 10 systematic subsamples were selected with 10 random starts, the procedure amounting to the selection of 10 arrays without replacement from a total of 30 arrays. In this particular situation, it happened that 6 consecutive arrays corresponding to 6 consecutive serial numbers were selected. It was known, however, that there was a high correlation between adjacent arrays. In the paper referred to, the situation was summarized as follows:

> "The statistician was then faced with a dilemma. If he used the sample, he could compute confidence limits and state that the population mean lay within those limits with the assurance that, if repeated samples were selected and analyzed the same way and a similar statement were made about the population mean on the basis of each analysis, the statements on the whole would be correct a high proportion of the time. But at the same time, it would be apparent to almost anyone who looked at the serial numbers selected that the particular sample in this instance was quite likely to be one of the exceptional cases where the statement would not be true. On the other hand, if the sample were discarded and a new sample were selected, the computation of a confidence interval would involve the difficulty that we do not know the risks involved in a procedure whereby we keep only samples that we like."

To avoid situations of this kind, it was suggested that a different kind of sampling procedure be employed. In particular it was suggested that the population to be sampled be divided into slices of equal size, and a random sample be selected independently from each slice. If such stratification were employed it would be extremely unlikely that 6 consecutive serial numbers would be drawn in every slice. It would still be theoretically possible, however. Thus, while it is desirable to employ sampling procedures such that samples we do

* Revision of paper presented at a luncheon meeting of the Chicago Chapter, American Statistical Association, November 15, 1956.

not like are highly improbable, even those procedures may not completely preclude the possibility of selecting a sample that is suspect. The suggested stratification is not a completely satisfactory solution to the problem.

A situation much like the one described may arise in experimental design. In an invited paper [7] before the Institute of Mathematical Statistics, Youden recently discussed the undesirability of using certain types of randomly chosen block arrangements of the experimental data. His suggestion is that such types be treated as unsatisfactory and excluded from the population of arrangements from which a random selection is to be made; but that inferences be drawn from a satisfactory arrangement, randomly selected, as if the selection were unrestricted. Other papers [2, 6] relating to restricted designs have been called to my attention by Kruskal.

The purpose of this paper is to suggest a simple alternative procedure which involves no assumption about the characteristics of the samples we exclude, and which may help to resolve the difficulty in certain favorable situations. As Kruskal has indicated (in private correspondence), the underlying principle can be extended to tests of hypotheses, where it can be applied to Type I or Type II errors, or both.

Thomas Hayton of the American Telephone and Telegraph Company has pointed out (verbally) that the problem of inadmissible samples is related to a problem in constructing tables of random numbers. These tables are commonly subjected to tests for randomness, and rejected or corrected if they fail to pass these tests. A question arises as to the possible effect of this censorship in computing confidence limits for sample statistics when a censored table is used to select the sample. The procedure proposed here can be extended to cover this situation, provided (a) the process used to generate the numbers is actually random, (b) an entire table of numbers would have been either accepted or rejected, but never patched up, and (c) the user of the table knows the joint probability that a table produced by the random process would not be used, and that a sample selected by employing the table would not be accepted as satisfactory. Condition (a), of course, is never completely realized.

Another related problem is that associated with wild values, or outliers, in experimental measurements where randomness is assumed, rather than deliberately introduced. Suppose, for example, that some measurements are assumed to be observations of a process in "statistical control"; that is, the measurements are assumed to behave as though they were random selections from a normally distributed universe. We might then want to test the most extreme observation for outliers, as by using the ratio of the most extreme deviation from the mean of the observations to the standard deviation of these observations. The statistic to be used, of course, will depend on the feared alternatives. The procedure proposed here could be applied in conjunction with a test of this kind, if we agree to discard all observations if the most extreme observation fails to pass the test, and if we know the probability of failure when the measuring process is actually in statistical control.

UNDERLYING THEORY

Let S denote the class of all possible samples that can result from some particular sampling procedure. Let this class be divided into two mutually exclu-

sive subclasses, A and N, such that every possible sample in the class S belongs to either A or N, but no sample belongs to both. For our purpose here, A denotes the subclass of all admissible samples, and N denotes the subclass of all samples that are not admissible, according to some specified criterion. This criterion for admissibility may be anything we choose, as long as it clearly divides the entire class S into the two sub-classes A and N.

Consider a fixed rule of inference, applicable to each possible sample, and let α_A and α_N denote the proportions of samples in subclasses A and N, respectively, for which this rule will lead to an incorrect inference. Then the probability for the entire class S that a randomly selected sample will lead to an incorrect inference is

$$\alpha = p_N\alpha_N + (1 - p_N)\alpha_A, \tag{1}$$

where p_N is the proportion of the samples in class S that belong to subclass N.

The problem considered here is essentially this: If we know α, what can we say about α_A?

Two possible answers to this question suggest themselves:

$$\alpha_A \leqq \alpha, \qquad \text{if } \alpha_A \leqq \alpha_N. \tag{2}$$

$$\alpha_A \leqq \frac{\alpha}{1 - p_N}, \qquad \text{for every } p_N \text{ and } \alpha_N. \tag{3}$$

These inequalities follow from the fact that p_N and α_N in equation (1) are necessarily nonnegative numbers.

YOUDEN'S PROCEDURE

Youden's suggestion appears to be an application of inequality (2). From experience with various types of experiments, he is convinced that for certain randomized arrangements in block designs the conditional risk of drawing an incorrect inference from experimental data is greater than for other arrangements. Hence, in making the random selection he excludes or discards the arrangements associated with the greater risk.

Suppose we designate an arrangement that is excluded an inadmissible arrangement. Then in our notation, p_N is the proportion of the total number of possible arrangements that are inadmissible, α_N is the proportion of inadmissible arrangements that will lead to an incorrect inference, α_A is the proportion of admissible arrangements that will lead to an incorrect inference, and α is the proportion of the entire class of both admissible and inadmissible arrangements that will lead to an incorrect inference. In most situations, α is the probability that appears in the published tables. In effect, what Youden says is the following: For a subclass N of possible arrangements that I shall designate as inadmissible, and a subclass A that excludes subclass N but includes all other possible arrangements, I am sure that $\alpha_N > \alpha_A$. I shall therefore restrict my random selection to subclass A. In drawing inferences, I shall rely on inequality (2), and take α_A as approximately equal to the α in the published tables.

In the situation described in the second paragraph of this paper, the question of whether or not to keep the sample with 6 consecutive arrays among the 10 arrays selected actually arose because it was believed that the risk of drawing an incorrect inference from samples belonging to the subclass that includes 6 or more consecutive arrays was greater than the probability in the published tables. In other words, it was believed that $\alpha_N > \alpha$, where α_N is the risk of drawing an incorrect inference from samples of size 10 with 6 or more consecutive serial numbers, and α is the probability in the published tables. The sample was therefore discarded, and a new sample was selected, largely because of this belief.

When the subclass of inadmissible samples is well defined in advance of the actual selection, it will make no difference whether inadmissible samples are excluded before the actual selection, or whether they are discarded after the fact. In either case, the exclusion can be defended when the condition following inequality (2) is actually true.

With respect to the question as to whether this condition is actually true, however, a statistician as such will often have no opinion. He may have to rely on the experience and judgment of someone familiar with the field in which the sample study or experiment is to be conducted. On the other hand, the statistician is frequently consulted in designing and interpreting samples or experiments for the very reason that his conclusions are presumed to be objective, and not influenced by a prior opinion, right or wrong, about the subject matter. The statistician is therefore not likely to be completely satisfied with inequality (2) in assessing the risk of an incorrect inference.

THE PROPOSED PROCEDURE

If α_A is known or can be conveniently estimated from the sample observations (as is the case for certain restricted designs) the statistician's course is clear. But there are cases in which α (or its maximum) can be estimated for the unrestricted design, while α_A is difficult to estimate. For these cases we suggest a procedure based on the general inequality (3).

Care must be exercised in applying this procedure. Thus, in the situation that was described in the second paragraph of this paper, the final decision to discard the sample with 6 consecutive arrays was made after computation showed there was less than 1 chance in 100 of getting 6 or more consecutive serial numbers in a sample of size 10 that is randomly selected without replacement from the serial numbers 1 to 30, inclusive. Discarding the sample and selecting another could therefore be justified (*rationalized* may be a better word) on the ground that rejecting samples of a type that would rarely be selected will not have much effect on the proportion of incorrect inferences in repeated samples. While the actual decision to discard was probably correct, the reason behind it is not completely satisfactory because the subclass of inadmissible samples was not well defined in advance. There may be some question as to just what samples would have been regarded as inadmissible, and therefore as to whether the probability of selecting an inadmissible sample was actually negligible.

The procedure we propose, under ideal conditions, would be as follows:

1. In designing a random sampling plan or experiment for the purpose of calculating a confidence interval based on the probability distribution of some sample statistic, make any necessary change in the usual procedure so that the probability of selecting an inadmissible sample or design arrangement will be small.

2. As a part of the sample design, proceed in the usual way to decide upon the maximum permissible risk of drawing an incorrect inference from the sample data or experimental results. Designate this risk as the value of α_A.

3. Before the actual selection of the sample items or of the design arrangement is made, decide precisely which samples or arrangements are to be regarded as inadmissible. The definition should be so chosen that any sample or arrangement will be classified as inadmissible if, in the judgment of the statistician or the research worker for whom the design is made, the risk of drawing an incorrect inference from such sample or arrangement, if selected, will exceed α_A.

4. Compute p_N, the probability of selecting an inadmissible sample or design arrangement. When use is made of a published table of random numbers, the value of $(1-p_N)$ should be the a priori probability that neither of two events will occur: (1) the random process used to generate the numbers will produce a table that would not be published, and (2) the sample or design arrangement resulting from use of a published table will not be considered to be admissible. Subtracting the sum of the two probabilities for these two events from one (thus treating the events as mutually exclusive) will result in a conservative value (lower bound) for $(1-p_N)$.

It should be understood that questions as to whether a random number table has actually been rejected before one has been found suitable for publication, and as to whether a sample has previously been found inadmissible, are not pertinent to the evaluation of p_N. The pertinent questions have to do with the a priori probabilities that a table will be rejected and that a sample will be classified as inadmissible.

5. Select the sample or design arrangement by the random method incorporated in the design, and keep and use it unless it is one of those previously defined as inadmissible. If it is inadmissible, repeat the selection until an admissible sample or design arrangement has been selected. Keep and use the first such admissible sample arrangement.

6. In calculating the confidence interval on the basis of the probability distribution of the sample statistic that was chosen for this purpose, proceed as if the maximum permissible risk of drawing an incorrect inference were

$$\alpha_A(1 - p_N), \tag{4}$$

rather than α_A.

ILLUSTRATIVE EXAMPLE

For illustrative purposes, consider the following situation. We wish to select a sample of 40 from a population of size 1000 that is serially numbered from 1

to 1000. The measurements of the property in which we are interested are known to have a symmetrical, unimodal distribution resembling the normal distribution; but the mean and variance of the distribution are unknown. Our purpose in selecting the sample is to compute 90 per cent confidence limits for the population mean. The population is known to be fairly homogeneous, and the most feasible and economical sampling procedure appears to be simple random sampling without replacement. The t-distribution seems suitable for computing the desired confidence limits. A table of random numbers is to be used in selecting the sample. The probability that the process used in generating these numbers would result in a table not regarded as satisfactory is deemed to be negligible. Adjacent individuals in the population are somewhat correlated in a positive direction, however, so that the inclusion of two or more consecutively numbered individuals in the sample seems undesirable. Aside from this, we wish to place no restrictions on the sample selection.

For this situation, the steps in the procedure we propose might turn out to be as follows:

1. We decide that unless the value of p_N is too large, as subsequently determined, we shall follow the sampling procedure that appears to be most feasible and economical: random sampling without replacement.

2. Since we wish to compute 90 per cent confidence limits, we set

$$\alpha_A = 1 - .90 = .10. \tag{5}$$

3. Since the selection of two or more consecutively numbered individuals seems undesirable, we define samples that include such individuals to be inadmissible, and all other samples to be admissible.

4. We compute p_N, the probability of selecting an inadmissible sample. This can theoretically be done when the definition of an inadmissible sample or arrangement depends entirely on certain combinations of sample serial numbers or letter sequences. For example, if we select a sample of size m by randomly selecting m numbers without replacement from the serial numbers $1, 2, \cdots, M$, and if we define an inadmissible sample as one in which the numerical difference between any two sample serial numbers is less than some integer d, then the value of $(1-p_N)$ is

$$\binom{M-(m-1)(d-1)}{m} \Big/ \binom{M}{m} = \frac{[M-(m-1)(d-1)]!(M-m)!}{[M-(m-1)(d-1)-m]!M!}. \tag{6}$$

If we define an inadmissible sample to be one in which there is at least one run of consecutive sample serial numbers larger than r, the value of $(1-p_N)$ is

$$\sum_{i=0}^{c} (-1)^i \binom{M-m+1}{i} \binom{M-i(r+1)}{m-i(r+1)} \Big/ \binom{M}{m}, \tag{7}$$

where c is the integer that satisfies the inequality

$$c \leqq \frac{m}{r+1} < c+1. \tag{8}$$

The derivation of these formulas is outlined in another paper [4]. In this example, formula (6) leads to

$$1 - p_N = .20, \text{ approximately.} \tag{9}$$

5. We next select 40 sample serial numbers without replacement from the numbers 1 to 1000, repeating the selection if necessary until we obtain a set of 40 numbers of which no two are consecutive. If the first 5 random numbers we select are 265, 482, 747, 395, and 481, for example, we discard all five numbers because 481 and 482 are consecutive, and start the selection over again, repeating the procedure as often as necessary to obtain a properly selected set of 40 nonconsecutive serial numbers. These numbers will designate the individuals to be included in the sample.

6. We compute the confidence limits

$$\bar{X} - t_\alpha s/\sqrt{40}, \tag{10}$$

$$\bar{X} + t_\alpha s/\sqrt{40}, \tag{11}$$

where, for the sample measurements $X_1, X_2, \cdots, X_{40}$,

$$\bar{X} = \frac{1}{40}\sum_{i=1}^{40} X_i, \tag{12}$$

$$s^2 = \frac{1}{39}\sum_{i=1}^{40} (X_i - \bar{X})^2, \tag{13}$$

$$\alpha = \alpha_A(1 - p_N) = (.10)(.20) = .02, \tag{14}$$

and t_α is the value of t, for 39 degrees of freedom, such that the probability that t will not exceed this value is $1-\frac{1}{2}$ of .02, or .99.

By interpolating in a table of the t-distribution (e.g., [5], p. 134) for 39 degrees of freedom, we find the value of t such that .99 is the probability that t will not exceed this value to be

$$t = 2.429, \text{ approximately.} \tag{15}$$

Hence, if $\bar{X} = 10$ and $s^2 = 9$, for example, the required confidence limits will be approximately equal to

$$10 - 2.429\sqrt{9/40} \doteq 8.85 \tag{16}$$

$$10 + 2.429\sqrt{9/40} \doteq 11.15. \tag{17}$$

ADVANTAGES AND DISADVANTAGES OF THE PROPOSED PROCEDURE

In designing sample surveys, it sometimes happens that the people for whom the survey is designed will raise the objection that the sample is "not representative." When this remark is made, I like to quote a remark by W. E. Deming [1] to the effect that "what we select are not representative samples, but probability samples." With this introduction, an explanation of what we mean by *probability sample* and *confidence limits* is usually sufficient to satisfy the skeptics, or at least, to silence the objections.

On the other hand, we try to use some common sense. And common sense tells us that while we can safely accept odds of better than 1 to 99 that the mean of a population will lie outside the computed 99 per cent confidence limits,

for example, *when the wager is made before the sample is selected,* we should be very foolish to accept the same odds if offered to us by an expert in the field *after he has been informed as to which individuals appear in the sample.*

The principal advantage of the procedure based on inequality (3) is that it enables us, to some extent, to use a little common sense in reconciling this difference between a priori and a posteriori probabilities. It takes into account the fact that information as to the particular subclass to which the sample belongs may be relevant in judging whether a rule of inference is likely to lead to an incorrect decision. It provides a mechanism for dealing with those specific situations where the serial numbers selected provide such information.

What we should do, it appears, is discuss the situation with our client, and decide *before the sample is selected* just what combinations of sample individuals will be regarded as "unsatisfactory." Such samples will be classified as inadmissible. In most cases, perhaps, the classification will have to be made on the basis of a maximum number of consecutive serial numbers, or on a minimum interval between serial numbers. While a sample of individuals from just one or two of several groups of geographical locations might also be classified as inadmissible, it would seem better to remove this difficulty by stratifying the sample if this is feasible. The discussion with the client should lead to an agreement that any sample not classified as inadmissible will, if selected, be retained as satisfactory.

The second advantage of the procedure based on inequality (3) is that it avoids any question as to the possible effect of bias in the estimates of the mean and variance, due to excluding inadmissible samples, on the computed confidence limits. If the quantity α is the true probability that the mean of the population will lie outside the confidence limits computed by some fixed rule, if applied to repeated samples of a given class from a given population, then the quantity $\alpha/(1-p_N)$ is an upper bound to the probability that the mean of the population will lie outside confidence limits computed by applying the same rule to an admissible subclass of samples. The statistician thus has an objective method of computing confidence limits from a sample that is classified as admissible.

A disadvantage of the proposed procedure is that computing the probability of selecting an inadmissible sample may be difficult or impossible. To illustrate, in sampling from a discrete universe, it might be desirable to exclude samples where all sample measurements happen to be alike, if it were known that the measurements in the universe were not all alike. But the probability of selecting a sample of that kind may be completely unknown. Even where the subclass of inadmissible samples is based entirely on combinations of serial numbers, the computation of the probability may be long and difficult.

Another disadvantage of the proposed procedure is that a confidence interval computed on the basis of inequality (3) will be wider than one computed by the methods now in use; or, the samples will have to be larger to obtain a confidence interval of a specified width. The reason is that the proposed procedure takes account of the possibility that the samples classified as inadmissible, presumably because the rule of inference is more likely to lead to an incorrect conclusion than for other samples, may actually be samples where the rule always leads to the correct conclusion. Thus, the proposed procedure takes the

most pessimistic view that is possible with respect to the judgment of the statistician and his client in choosing those sample combinations likely to lead to trouble. It is something like charging a higher insurance premium (than would be appropriate for the whole population) to those who pass a medical examination because some individuals are rejected on the basis of the examination.

The increased sample size, or the wider confidence interval, is the price that seems necessary to pay, however, if the statistician is to compute a confidence interval in which he can actually have confidence on the sole basis of his specialized knowledge. This price may be exorbitant in cases where there is a fairly large proportion of possible samples with respect to which the client would be more or less dubious about the probabilities computed by applying a rule of inference.

On the other hand, if an expert in the field is willing to employ inequality (2) and state that, in his opinion, the mean of a sampled population lies within narrower limits than those computed by applying inequality (3), there would seem to be no objection to including his statement in the statistician's report. The client could then take both sets of confidence limits into account in making his decision.

Perhaps the most serious drawback to general use of the proposed procedure is the fact that it requires the subclass of inadmissible samples to be precisely defined in advance. It is likely to be difficult or impossible to secure a good definition of this kind. Ideally, we should know enough about the universe we are sampling, about the estimation of limits of sampling variability, and about balancing precision against cost and administrative feasibility to select a sample design which makes "unsatisfactory" samples highly improbable. In real life, however, we must be prepared to find, after the sample has been selected, that we have overlooked something so obvious that we cannot defend our design without further restriction.

Nevertheless, keeping the ideal procedure before us as a goal may help to minimize the number of unsatisfactory samples. The primary purpose of this paper is to help attain that objective.

REFERENCES

[1] Deming, W. E., informal remarks in A Seminar on Scientific Sampling at the New Jersey Bell Telephone Company, April 29 to June 16, 1956.

[2] Grundy, P. M., and Healy, M. J. R., "Restricted randomization and quasi-Latin squares," *Journal of the Royal Statistical Society*, Series B, 12 (1950), 286–91.

[3] Jones, Howard L., "Investigating the properties of a sample mean by employing random subsample means," *Journal of the American Statistical Association*, 51 (1956), 54–83.

[4] Jones, Howard L., "Maximum runs and minimum intervals in sampling integers without replacement," submitted for publication.

[5] Pearson, E. S., and Hartley, H. O., *Biometrika Tables for Statisticians*, Vol. I, Cambridge University Press, 1954.

[6] Yates, F., "Bases logiques de la planification des expériences," *Annales de l'Institute Henri Poincaré*, 12 (1951), 97–112.

[7] Youden, W. J., "Experimental designs for industrial research," unpublished paper presented at the 72nd meeting of the Institute of Mathematical Statistics, Detroit, Michigan, September 8, 1956.

THE PRECISION OF UNBIASED RATIO-TYPE ESTIMATORS*

LEO A. GOODMAN
University of Chicago
AND
H. O. HARTLEY
Iowa State College

TABLE OF CONTENTS

1. INTRODUCTION 491
2. THE BIAS IN RATIO ESTIMATORS 492
3. FORMULAS FOR BIAS AND ITS ELIMINATION 493
4. THE PRECISION OF THE BIASED AND UNBIASED RATIO ESTIMATORS 493
4.1. The Comparison of the Estimators y' and $\hat{y}$ 494
4.2. The Comparison of the Estimators y' and $\bar{y}$ 495
5. THE ESTIMATION OF THE VARIANCE OF y' 497
6. SOME EXAMPLES 500
7. SOME SPECIAL CASES 502
8. ON SOME REGRESSION ESTIMATORS 506
9. REFERENCES 507

When each unit in a population of N units consists of an x and a y measurement, where the population mean $\bar{X}$ of the x's is known, then the population mean $\bar{Y}$ of the y's is frequently estimated by drawing a random sample of n units and using one of the customary biased ratio-type estimators, $\bar{y}$ (based on the ratio of the y sample mean and the x sample mean), or $\hat{y}$ (based on the sample mean of the ratios of the y and x measurements for each unit). The precision of these estimators is compared with an exact formula derived here (when $N \gg n$) for the variance of an unbiased ratio-type estimator y', a modification of $\bar{y}$ and $\hat{y}$, introduced in [5]. Conditions pertaining to the population moments are obtained under which Var $y' <$ Var $\hat{y}$, and it is also shown that Var $y' <$ Var $\bar{y}$ (when $n \to \infty$) if and only if the slope of the population regression line of y on x is closer to the population mean of the ratios of the y and x measurements for each unit than to $\bar{Y}/\bar{X}$. The estimation of the variances of the three ratio-type estimators is discussed, and some numerical examples are presented to illustrate the comparison of the estimators. Certain special populations are examined in detail, and the relation between regression and ratio-type estimators is also discussed.

1. INTRODUCTION

THIS paper is concerned with a modification of the usual ratio estimators that leads to an unbiased estimator. This property of unbiasedness is important in certain sample surveys involving a large number of strata. The precision of the unbiased ratio estimator is compared with the usual ratio estimators. Both the situation where the sample size in each stratum is small and the situation where the sample size is large will be considered in the present paper. In most of this paper, no assumptions are made about the population distributions (e.g., no assumptions concerning normality appear here), and the results obtained do not depend on such assumptions.

* A part of this work was carried out under the sponsorship of the Statistics Branch, Office of Naval Research.

D. S. Robson, in a recent paper [9], refers to the unbiased ratio-type estimator and the formula for its variance presented in the present paper (actually, he refers to an earlier unpublished version of the present paper), and he computes the exact variance formula using multivariate polykays. If the size N of the population is large relative to the sample size n, the variance formula presented in the present paper (which is "exact" when $N \gg n$) is recommended, since it is simpler to compute than the exact formula in [9]. However, the latter formula is exact in the situation where N may not be large relative to n, and should be used in that case if exact formulas are important.

The existence of bias in the standard ratio estimators obtained from random samples has led to the exploration of the possibilities of modifying the sampling scheme so that the ratio estimator would become unbiased for the modified sampling scheme (see [8] and the articles referred to therein). In the present paper, the sampling is not modified (i.e., the case of a simple random sample drawn from a finite population is considered here), but the usual ratio estimators are modified so that a ratio-type estimator is obtained that is unbiased for the usual simple random sampling scheme.

2. THE BIAS IN RATIO ESTIMATORS

Ratio estimators are frequently employed in sample surveys when estimating the population mean $\overline{Y}$ of a variable y with the help of the known population mean $\overline{X}$ of a variable x that is positively correlated with y. In the case of a simple random sample of n pairs (x_i, y_i) drawn from a population of N pairs, various ratio estimators of $\overline{Y}$ can be formed. The following two estimators, which are the usual ratio estimators, will be considered here:

$$\bar{y} = \overline{X}\bar{y}/\bar{x} \quad \text{and} \quad \bar{y} = \overline{X}\bar{r}, \tag{1}$$

where $\bar{y}$, $\bar{x}$ and $\bar{r}$ are the arithmetic means of the sample of y_i, x_i and $r_i = y_i/x_i$ respectively (see p. 111 and p. 124 in [1]). In this paper we shall assume that $x_i > 0$, and that $\overline{X}$ is known. (In some surveys, when some $x_i = 0$, the use of larger sampling units, so that the new x_i being averages of the original x_i are positive, can be recommended for certain purposes in some of these situations.)

One well-known defect of these estimators is the fact that they are usually biased by amounts for which only approximate formulas are offered in the literature. Often in practical surveys, the bias is a small fraction of the standard deviation of the estimate and can be neglected. There is, however, an important class of survey designs in which this bias may become considerable. If a small sample of (x_i, y_i) is drawn from each of a large number (k) of strata, then it will often be found that the bias in the estimate of each stratum total is of the same sign, and the bias in the estimate of the population total will be approximately k times that for an individual stratum total while the standard deviation only multiplies by $\sqrt{k}$ (see e.g., p. 130 in [1]). Therefore, the mean squared error of the estimate of the population total will, in this case, be of the order of magnitude k^2; while, if the estimate of each stratum total is unbiased, the order of magnitude will be k. It is in these situations that the use of unbiased ratio estimators is of great advantage.

3. FORMULAS FOR BIAS AND ITS ELIMINATION

In a note [5], simple formulas for the bias in $\hat{\bar{y}}$ and $\bar{y}$ were given, which led to the construction of an unbiased ratio estimator. These results may be restated as follows. By definition of the population covariance, we have the equations

$$E(\hat{\bar{y}}) = \overline{Y} - \mathrm{Cov}\left(\frac{y}{x}, x\right) \tag{2}$$

$$E(\bar{y}) = \overline{Y} - \mathrm{Cov}\left(\frac{\bar{y}}{\bar{x}}, \bar{x}\right) \tag{3}$$

where the symbol $E(z)$ denotes the expectation for the population of all possible sample values of z computed for samples of size n. (This incorporates as a special case ($n=1$) the usual expectation over the N values of the finite population.) The equations (2) and (3) yield convenient and exact expressions for the bias. An immediate consequence of (2) and (3) are upper bounds for the biases; viz.,

$$|\text{Bias in } \hat{\bar{y}}| \leq \sigma_{y/x} \cdot \sigma_x, \text{ and } |\text{Bias in } \bar{y}| \leq \sigma_{\bar{y}/\bar{x}} \cdot \sigma_{\bar{x}}. \tag{4}$$

These upper bounds are proved (but only as large sample approximations) in some textbooks.

Let us now consider the elimination of the bias in $\hat{\bar{y}}$. Clearly an unbiased estimate of Cov $(y/x, x)$ is given by

$$\frac{(N-1)}{N(n-1)} \sum_{i=1}^{n} (r_i - \bar{r})(x_i - \bar{x}) = \frac{(N-1)n}{N(n-1)} (\bar{y} - \bar{r}\bar{x}).$$

Hence,

$$y' = \overline{X}\bar{r} + \frac{(N-1)n}{N(n-1)} (\bar{y} - \bar{r}\bar{x}) \tag{5}$$

is an unbiased estimator of $\overline{Y}$.

4. THE PRECISION OF THE BIASED AND UNBIASED RATIO ESTIMATORS

The usual variance formulas for the customary biased ratio estimators $\hat{\bar{y}}$ and $\bar{y}$, which are available in the literature, are approximate. An approximation for the variance of $\bar{y}$ is usually derived as a low order expansion in powers of $\delta\bar{x} = (\bar{x} - \overline{X})/\overline{X}$ or by an equivalent approximation (see e.g. [1], p. 115). The approximation for the variance of $\hat{\bar{y}}$ is usually derived as a low order expansion in powers of $\delta x_i = (x_i - \overline{X})/\overline{X}$ or by equivalent approximations. With the former the $\delta\bar{x}$ are assumed small in absolute value, while with the latter this assumption is made for δx_i (e.g. [1], p. 117). Moreover, it is often assumed that N is large compared with $n (N \gg n)$. In the present paper, we make throughout the assumption that $x_i > 0$; i.e., that $-1 < \delta x_i$. This is the only assumption required for the validity of subsequent exact formulas. However, when discussing the

relative order of magnitude of certain terms in certain of our formulas, we base this discussion on one of two further assumptions; viz.,

(a) $$|\delta x_i| \ll 1$$

or

(b) $$|\delta x_i| \ll 1 \quad \text{and} \quad |\delta y_i| \ll 1,$$

where $\delta y_i = (y_i - \overline{Y})/\overline{Y}$. These latter assumptions, the reader will note, can be quite restrictive. Deriving an approximate variance formula for y' under assumption (b), we obtain

$$\text{Var } y' = n^{-1}\overline{Y}^2\{\cdot\sigma_y^2/\overline{Y}^2 + \sigma_x^2/\overline{X}^2 - 2 \text{ Cov } xy/\overline{X}\overline{Y}\} \tag{6}$$

provided $n \ll N$. This formula was obtained by retaining all terms up to and including cubic terms in δx_i and δy_i. Formula (6) will be identified as the familiar approximate formula for both Var $\hat{y}$ and Var $\bar{y}$ if cubic and higher order terms in the respective formulas are neglected.

4.1. *The comparison of the estimators y' and $\hat{y}$*

In order to compare the variances of the ratio estimators y' and $\hat{y}$, we develop exact variance formulas. It is convenient to derive such formulas in terms of the joint population distribution of the x_i and $r_i = y_i/x_i$ rather than that for the x_i and y_i. For the sake of simplicity, we shall continue to assume a large or an infinite population ($N \gg n$), although it is possible to derive the corresponding results for finite populations.

The estimate $\hat{y} = \overline{X}\bar{r}$ is biased by an amount $(\overline{R}\overline{X} - \overline{Y})$, where $\overline{R}$ is the mean for the population of the r_i values. Hence, the estimate will not be consistent unless $\overline{R} = \overline{Y}/\overline{X}$. The exact variance formula for $\hat{y}$ is

$$\text{Var } \hat{y} = \overline{X}^2\sigma_r^2/n, \tag{7}$$

We can simplify the formula for our unbiased estimator y' to

$$y' = \overline{X}\bar{r} + c, \tag{8}$$

where

$$c = (n-1)^{-1}\sum_{i=1}^{n}(x_i - \bar{x})(r_i - \bar{r}) \tag{9}$$

is an unbiased estimate of the population covariance between x and r. The exact variance for y' is therefore

$$\text{Var } y' = n^{-1}\overline{X}^2\sigma_r^2 + 2\overline{X} \text{ Cov } \bar{r}c + \text{Var } c. \tag{10}$$

The second term in (10), after some straightforward calculations, can be written as

$$2\overline{X}n^{-1}E[(\Delta x)(\Delta r)^2], \tag{11}$$

where $\Delta x = (x - \overline{X})$, $\Delta r = (r - \overline{R})$. For the third term in (10), we employ standard techniques of using the moments of bivariate sample cumulants (see [2] and [3]). From this we obtain

$$\text{Var } c = \frac{1}{n}\left\{E[(\Delta x)^2(\Delta r)^2] + \frac{\sigma_r^2\sigma_x^2}{n-1} - \frac{n-2}{n-1}C_{rx}^2\right\}, \tag{12}$$

where $\sigma_r^2 = E[(\Delta r)^2]$ and $\sigma_x^2 = E[(\Delta x)^2]$ are the population variances of r and x respectively, and $C_{rx} = E(\Delta r \Delta x)$ is the population covariance. Combining these terms, we have

$$n \operatorname{Var} y' = \overline{X}^2\sigma_r^2 + 2\overline{X}E[(\Delta x)(\Delta r)^2] + E[(\Delta x)^2(\Delta r)^2] + \frac{\sigma_r^2\sigma_x^2}{n-1} - \frac{(n-2)C_{rx}^2}{n-1} \tag{13}$$

which is the exact variance formula for y'.

In terms of δx_i and δy_i, the first term (which is n Var $\bar{y}$) in (13) is of 2nd order, the second (covariance term) is of third order and the 3rd, 4th, and 5th terms are of 4th order. Hence, if the coefficients of variation of x and y are small (assumption (b)), the 2nd term will be the important one when comparing Var y' and Var $\bar{y}$. This term will be negative if the array variance of r given x, diminishes with increasing x (see Section 6 for some data that illustrate this case). In this case, the estimate y' is to be preferred to $\bar{y}$.

In situations where the coefficients of variations are not smaller than say $\frac{1}{4}$, we may not be able to neglect entirely the 3rd, 4th and 5th term in (13). Noting that

$$\operatorname{Var}(\Delta r \Delta x) = E[(\Delta r)^2(\Delta x)^2] - C_{rx}^2, \tag{14}$$

and introducing the regression coefficient

$$\beta = E[(\Delta r)^2 \Delta x]/\operatorname{Var}(\Delta r \Delta x) \tag{15}$$

of Δr on the product $\Delta r \Delta x$, equation (13) may be rewritten in the form

$$n \operatorname{Var} y' = \overline{X}^2\sigma_r^2 + \operatorname{Var}(\Delta r \Delta x)\{2\overline{X}\beta + 1\} + \frac{1}{n-1}(\sigma_x^2\sigma_r^2 + C_{rx}^2). \tag{16}$$

The third term in this formula is not larger than $2\sigma_x^2\sigma_r^2/(n-1)$. Hence, this third term represents a fraction not larger than $2\sigma_x^2/\overline{X}^2(n-1)$ of the leading first term $\overline{X}^2\sigma_r^2$, and will usually be negligible. If this term is neglected, the criterion as to whether Var $y' <$ Var $\bar{y}$ or vice versa will depend on whether or not $\beta \leq -1/(2\overline{X})$. Thus, if the array variance of r diminishes with increasing x, this will indicate that $E[(\Delta x)(\Delta r)^2] < 0$; i.e., $\beta < 0$. If $\beta < 0$ and $\overline{X}$ is large when compared with $|\Delta x|$, then it will be found, in many cases, that $\beta \leq -1/(2\overline{X})$. A clearer indication as to whether the inequality Var $y' <$ Var $\bar{y}$ is satisfied can be obtained by computing unbiased variance estimates for both y' and $\bar{y}$ as indicated in Section 5.

In this section, we have indicated when Var $y' <$ Var $\bar{y}$. If it is, then y' is to be preferred to $\bar{y}$ since y' is unbiased and also has a smaller variance. However when Var $y' \geq$ Var $\bar{y}$, $\bar{y}$ is not necessarily to be preferred to y', since the bias of $\bar{y}$ may be large. In terms of δx_i the bias is of the first order, and it does not diminish as n increases. Hence, except in the case where $\overline{R} = \overline{Y}/\overline{X}$, the bias of $\bar{y}$ must also be considered when comparing this estimator with y'.

4.2 *The comparison of the estimators y' and $\bar{y}$*

For the comparison of the variances of $\bar{y}$ and y', we shall introduce moment-coefficients for both the joint distribution of y and x as well as x and r. Let us

first simplify the exact variance formula for y'. Since

$$\begin{aligned}\text{Var}\,(y - \overline{R}x) &= \text{Var}\,(\Delta x \Delta r + \overline{X}\Delta r) \\ &= \text{Var}\,(\Delta x \Delta r) + \overline{X}^2\sigma_r^2 + 2\overline{X}E[\Delta x(\Delta r)^2] \qquad (17) \\ &= \sigma_y^2 + \overline{R}^2\sigma_x^2 - 2\overline{R}C_{yx},\end{aligned}$$

equations (13) and (16) may be rewritten as

$$n\ \text{Var}\ y' = (\sigma_y^2 + \overline{R}^2\sigma_x^2 - 2\overline{R}C_{yx}) + (n-1)^{-1}(\sigma_r^2\sigma_x^2 + C_{rx}^2), \qquad (18)$$

which is a simple form for the exact variance of the unbiased ratio estimator y'.

Let us now consider the mean square error $E(\bar{y}-\overline{Y})^2$ of $\bar{y}$. We have

$$\text{M.S.E.} = E(\bar{y} - \overline{Y})^2 = \overline{Y}^2\overline{X}^2E\left\{\frac{\delta\bar{y} - \delta\bar{x}}{\bar{x}}\right\}^2 = \overline{Y}^2E\left\{\frac{\delta\bar{y} - \delta\bar{x}}{1 + \delta\bar{x}}\right\}^2 \qquad (19)$$

where $\delta\bar{y}=(\bar{y}-\overline{Y})/\overline{Y}$ and $\delta\bar{x}=(\bar{x}-\overline{X})/\overline{X}$. Hence, the mean square error is equal to the value, at argument $\theta=1$, of the function

$$f(\theta) = \overline{Y}^2E\left\{\frac{\delta\bar{y} - \delta\bar{x}}{1 + \theta\delta\bar{x}}\right\}^2 \qquad (20)$$

and may therefore be represented by its Taylor expansion, at $\theta=0$, with exact 3rd derivative remainder term. This expansion yields

$$\begin{aligned}\overline{Y}^{-2}(\text{M.S.E.}) &= E(\delta\bar{y} - \delta\bar{x})^2 - 2E\{(\delta\bar{y} - \delta\bar{x})^2\delta\bar{x}\} \\ &\quad + 3E\{(\delta\bar{y} - \delta\bar{x})^2(\delta\bar{x})^2\} - 4E\left\{\frac{(\delta\bar{y} - \delta\bar{x})^2(\delta\bar{x})^3}{(1 + \theta\delta\bar{x})^5}\right\}. \qquad (21)\end{aligned}$$

Since the 4th derivative $f^{(iv)}(\theta)$ is ≥ 0 for $0\leq\theta\leq 1$, $f^{(iii)}(\theta)$ is monotonically increasing and the third order remainder term must lie between the two values obtained when putting $\theta=0$ and $\theta=1$ respectively; i.e., between the limits

$$L_1 = -4E\{(\delta\bar{y} - \delta\bar{x})^2(\delta\bar{x})^3\}$$

and

$$L_2 = -4E\left\{\frac{(\delta\bar{y} - \delta\bar{x})^2(\delta\bar{x})^3}{(1 + \delta\bar{x})^5}\right\}. \qquad (22)$$

The expansion (21) proceeds in powers of $\delta\bar{x}$ the first term being of order 0 and the last of 3rd order. For large values of n, both $\delta\bar{x}$ and $\delta\bar{y}$ are of order $1/\sqrt{n}$. Therefore, the leading term in (21), which is of order n^{-1}, yields the asymptotic large sample formula for $\overline{Y}^{-2}$ (M.S.E.). Since the square of the bias $(\text{Cov}^2(\bar{y}/\bar{x}, \bar{x}))$ is clearly of order n^{-2}, the leading term in (21) also yields the large sample formula for $\overline{Y}^{-2}$ Var $\bar{y}$. Thus as $n\to\infty$,

$$\text{Var}\ \bar{y} \to \overline{Y}^2E(\delta\bar{y} - \delta\bar{x})^2 = n^{-1}(\sigma_y^2 + Q^2\sigma_x^2 - 2QC_{yx}) \qquad (23)$$

where $Q=\overline{Y}/\overline{X}$.

The corresponding large sample formula for Var y' is from (18) obtained as

$$\text{Var } y' \to n^{-1}(\sigma_y^2 + \overline{R}^2\sigma_x^2 - 2\overline{R}C_{yx}). \tag{24}$$

Hence, asymptotically for large n, we have

$$\text{Var } \bar{y} - \text{Var } y' = n^{-1}\sigma_x^2\{(Q - B)^2 - (\overline{R} - B)^2\} \tag{25}$$

where $B = C_{yx}/\sigma_x^2$ is the regression coefficient of y on x. Equation (25) shows that when n is large, Var $\bar{y}$ will be larger than Var y' if and only if the regression coefficient B is nearer to $\overline{R} = E(y/x)$ than to $Q = \overline{Y}/\overline{X}$. We also find that Var $\bar{y}$ will be larger than Var y' if and only if line $\overline{Y} + \overline{R}(x - \overline{X})$ "fits the scatter diagram of y on x" better than the line $\overline{Y} + Q(x - \overline{X}) = Qx$. In particular, the variances agree if $\overline{R} = Q$; i.e., if the ratio estimator $\hat{y}$ is unbiased. When $\overline{R} \neq Q$, the regression coefficient B will, in our experience, often be closer to Q than to $\overline{R}$, and in such cases the large sample formula for Var y' will be larger than that for Var $\bar{y}$.

When the sample is large, the difference between the variances of the two estimates can be estimated by

$$n^{-1}s_x^2\{(q - b)^2 - (\bar{r} - b)^2\}$$

where s_x^2 is the usual unbiased sample estimate of the variance of x, b is the usual estimate of the regression coefficient of y on x, and $q = \bar{y}/\bar{x}$. The variances of both $\bar{y}$ and y' can be estimated in a similar manner using the appropriate estimates of the terms in equations (23) and (24). The appropriate estimate of the terms in (23) can be obtained using a similar method to that given on p. 119 in [1]. Using this method, an estimate may also be obtained for (24).

The difference between the variances of $\bar{y}$ and y' can be written

$$\begin{aligned} \text{Var } \bar{y} - \text{Var } y' &\doteqdot n^{-1}\sigma_x^2(Q + \overline{R} - 2B)(Q - \overline{R}) \\ &\doteqdot n^{-1}\sigma_x^2 C_{xr}(Q + \overline{R} - 2B)/\overline{X}, \end{aligned} \tag{26}$$

which shows that, under assumption (a), this difference is of the order of magnitude $n^{-1}\delta_x^3$. When n is small, it does not seem possible to derive a simple criterion for the comparison of Var $\bar{y}$ and Var y' in the general case, but some special cases are considered in Section 7.

5. THE ESTIMATION OF THE VARIANCE OF y'

In this section, two types of formulas for variance estimation are developed. The first is directly derived from the formula for the variance (10) and arises in the orthodox way of replacing each term in this formula by an unbiased estimate from the sample values. We again assume an infinite population and make use of bivariate k-statistics, which are unbiased estimators of the corresponding population cumulants (see [2] and [3]). With the help of this theory, the following results can be assembled: Defining by

$$s_r^2 = (n - 1)^{-1}\sum (r_i - \bar{r})^2, s_x^2 = (n - 1)^{-1}\sum (x_i - \bar{x})^2,$$
$$c = (n - 1)^{-1}\sum (x_i - \bar{x})(r_i - \bar{r})$$

the sample variances and covariance, and by σ_r^2, σ_x^2 and C_{rx} the corresponding population parameters, we have

$$E(s_r^2) = \sigma_r^2$$

$$E(c^2) = \frac{1}{n} K_{22} + \frac{1}{n-1} \sigma_r^2\sigma_x^2 + \frac{n}{n-1} C_{rx}^2 \tag{27}$$

$$E(s_r^2 s_x^2) = \frac{1}{n} K_{22} + \sigma_r^2\sigma_x^2 + \frac{2}{n-1} C_{rx}^2, \tag{28}$$

where the cumulant K_{22} is defined as

$$K_{22} = E[(\Delta x)^2(\Delta r)^2] - \sigma_x^2\sigma_r^2 - 2C_{rx}^2. \tag{29}$$

An unbiased estimate of the cumulant K_{22} is the k statistic k_{22} (see p. 281 in [6]), where

$$k_{22} = \frac{n}{(n-1)(n-2)(n-3)} \left\{ (n+1)S_{22} - \frac{2(n+1)}{n} S_{21}S_{01} - \frac{2(n+1)}{n} S_{12}S_{10} \right.$$

$$- \frac{(n-1)}{n} S_{20}S_{02} - \frac{2(n-1)}{n} S_{11}^2 - \frac{8}{n} S_{11}S_{10}S_{01} \tag{30}$$

$$\left. + \frac{2}{n} S_{20}S_{01}^2 + \frac{2}{n} S_{02}S_{10}^2 - \frac{6}{10} S_{10}^2 S_{01}^2 \right\}$$

and

$$S_{ij} = \sum_{i=1}^{n} x_i^i r_i^j$$

We now note that

$$E(k_{22}) = K_{22}, \tag{31}$$

and introducing

$$c' = \frac{1}{(n-1)} \sum_{i=1}^{n} (r_i - \bar{r})^2 (x_i - \bar{x}), \tag{32}$$

we find that

$$E(c') = \frac{n-2}{n} E[(\Delta r)^2 \Delta x]. \tag{33}$$

Equations (27), (28), (29) and (31) permit the calculation of unbiased estimators of $E[(\Delta x)^2(\Delta r)^2]$, $\sigma_x^2\sigma_r^2$ and C_{rx}^2, and hence of Var c. Equations (32) and (33) yield an unbiased estimator of $E[(\Delta r)^2 \Delta x]$. Combining these results, we have, for samples of size $n \geq 3$, formulas for unbiased estimators, $v(\bar{y})$ and $v(y')$, of the variances $V(\bar{y})$ and $V(y')$ respectively; viz.,

$$nv(\bar{y}) = \bar{X}^2 s_r^2, \tag{34}$$

and

$$v(y') = n^{-1}\overline{X}^2 s_r^2 + 2\overline{X}c'/(n-2) + \frac{(n-1)s_r^2 s_x^2 + (n-3)c^2 + \left(1-\frac{2}{n}\right)(n-1)k_{22}}{n^2-n-2} \tag{35}$$

where the computational formulas for c, c' and k_{22} are as follows:

$$\begin{aligned} n(n-1)c &= n\sum y_i - (\sum x_i)(\sum r_i), \\ (n-1)c' &= \sum y_i r_i - 2\sum y_i + \bar{r}^2 \sum x_i - (n-1)\bar{x}s_r^2 \end{aligned} \tag{36}$$

and k_{22} is computed from (30) remembering that in our case

$$\begin{aligned} &S_{22} = \sum y_i^2,\ S_{21} = \sum y_i x_i,\ S_{12} = \sum y_i r_i,\ S_{20} = \sum x_i^2,\ S_{02} = \sum r_i^2 \\ &S_{11} = \sum y_i,\ \ S_{10} = \sum x_i,\ \ S_{01} = \sum r_i \end{aligned} \tag{37}$$

The formula for $v(y')$ is computationally cumbersome, and it is therefore of interest to develop a simpler method of variance estimation. This second estimate of variance, although again unbiased, will usually be affected by a larger sampling error. Moreover, with this method, it is convenient to modify the sampling scheme as follows: Suppose that we are sampling from a population of N pairs (x, y), and that $n=2m$ is even and $m \geq 2$. Draw a first random sample of m pairs (x_i, y_i) without replacement. Next replace the whole sample and draw a second sample of m pairs (x_i, y_i) without replacement. This method of sampling makes the two samples of size m independent. (For the procedure of splitting at random the sample of size n into two samples of size m, this will not be the case.) If $n \ll N$, the two samples will usually have no units in common. Very occasionally, however, the samples may have one or more units in common, and we specify that if the second sample is completely identical with the first it must be rejected and a new sample drawn. We may think of the above scheme as follows: From the original population of N pairs of (x, y) form a superpopulation of the $\binom{N}{m}$ possible samples of size m, and draw two of these samples without replacement. If we denote by $\bar{x}_1$, $\bar{y}_1$, $\bar{r}_1$ and $\bar{x}_2$, $\bar{y}_2$ and $\bar{r}_2$ the sample means of the x_i, y_i and r_i for the two samples, we may form the two unbiased ratio estimators

$$y_1' = \overline{X}\bar{r}_1 + \left(\frac{N-1}{N}\right)\left(\frac{m}{m-1}\right)(\bar{y}_1 - \bar{r}_1\bar{x}_1)$$

$$y_2' = \overline{X}\bar{r}_2 + \left(\frac{N-1}{N}\right)\left(\frac{m}{m-1}\right)(\bar{y}_2 - \bar{r}_2\bar{x}_2).$$

Hence, $\bar{y}' = \frac{1}{2}(y_1' + y_2')$ is an unbiased estimate of $\overline{Y}$, and an unbiased estimate of the variance of $\bar{y}'$ is given by

$$\frac{1}{4}(y_1' - y_2')^2 \left\{ \frac{\binom{N}{m} - 2}{\binom{N}{m}} \right\}.$$

The finite sampling corrections $[(N-1)/N]$ and certainly

$$\left[\binom{N}{m} - 2\right] \Big/ \binom{N}{m}$$

can usually be replaced by 1. The unbiased estimate of the variance is based only on a "single degree of freedom." If more "degrees of freedom" are desired, K samples of size n/k could, of course, be drawn ($K \geq 2$). The disadvantage of a "single degree of freedom" (i.e., $K=2$) is eliminated to a certain extent in stratified sampling where the number of strata are large.

This method of estimation may be illustrated by the following example where the population size $N=400$ and $\overline{X}=2$. Two samples of size $m=2$ are drawn as follows:

1st Sample			2nd Sample		
x	y	r	x	y	r
1	3	3	4	8	2
2	6	3	1	2	2
$\bar{x}_1=1.5$	$\bar{y}_1=4.5$	$\bar{r}_1=3$	$\bar{x}_2=2.5$	$\bar{y}_2=5$	$\bar{r}_2=2$

Ignoring finite sampling corrections, we have that $y_1'=6$, $y_2'=4$, $\bar{y}'=5$, and the estimated variance of $\bar{y}'$ is $\frac{1}{4}(6-4)^2=1$.

6. SOME EXAMPLES

The comparison of the variance of the unbiased estimator y' and the biased estimator $\bar{y}$ is critically dependent on whether the array variance of r *decreases* with increasing x or not in the (x, r) scatter diagram. In Fig. 501, we have set out an example of such a scatter diagram. Fig. 501 deals with a sample of 49 large U. S. cities (see p. 113 in [1]) drawn from a larger population of $N=196$ cities. For this figure, x_i = population size in 1920, y_i = population size in 1930 and $r_i=y_i/x_i$ = relative population size. Fig. 501 indicates that the array variance of r diminishes with increasing x; i.e., the variation in *relative* size is more stable for the larger cities. Although little is proved by this isolated example, it is clear that many situations in which ratio estimators are appropriate are in fact represented by models in which the conditional variance of r given x is decreasing with x. Among these models are the following:

(a) x = size of population
y = number of population having a certain character (say sickness)
$r=y/x$ rate (say sickness rate).

If the binomial distribution applies, Var $(r|x)=x^{-1}r_x(1-r_x)$ where r_x is the array mean of r given x. This variance will normally decrease with increasing x.

(b) The array variance of y given x, Var $(y|x)$, is constant. In this case, Var $(r|x)$ will be proportional to $1/x^2$.

(c) Var $(y|x)$ is proportional to x. In this case, Var $(r|x)$ is proportional to $1/x$.

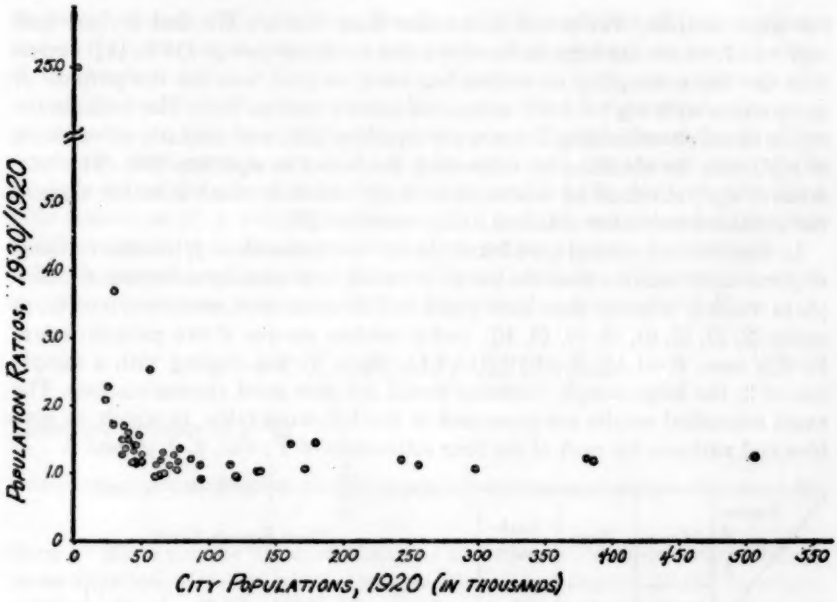


FIG. 501 U. S. Cities: Relative Population Increase, 1920–1930.

The following model is an illustration of a case where the conditional variance of r given x is constant:

(d) The case where Var $(y|x)$ is proportional to x^2 is sometimes one in which r is independent of x and the ratio estimator $\hat{y}$ is unbiased. In this case Var $(r|x)$ is constant, and Var y' will usually be larger than Var $\hat{y}$.

We shall discuss some of these models in more detail in the following section.

In order to illustrate the methods given in the present paper, we have used the sample data for the 49 cities listed in [1] and plotted in Fig. 501. For this population of $N=196$ cities, *both* $\overline{Y}$ as well as $\overline{X}$ are known. We have $\overline{Y}=149.8$ and $\overline{X}=116.9$. Let us first compare $\hat{y}$ and y'. Using the known value of $\overline{X}$, estimates of $\overline{Y}$ are

$$\hat{y} = 214.1 \qquad y' = 152.1.$$

(One reason for the poor estimate obtained from $\hat{y}$ is the very unusual growth of a city from 2,000 to 50,000 for which $r=25$ is "off the graph" in Fig. 1GH.) If N is taken as ∞ and if formulas (34) and (35) are used to estimate the variances of $\hat{y}$ and y' respectively, from the sample of 49 cities, we obtain:

$$v(\hat{y}) = 3248 \qquad v(y') = 113.7.$$

Hence, for this kind of data, the estimate y' is better than $\hat{y}$.

We shall now compare the ratio estimates y' and $\bar{y}$. In Section 4.2, we noted that for large samples Var $\bar{y}$ will be larger than Var y' if and only if the regression coefficient B of y on x is nearer to $\overline{R}$ than to $Q=\overline{Y}/\overline{X}$. For the data plotted in Fig. 501, we find that $\bar{y}=144.9$, and $B=1.2$, $\overline{R}=1.8$, $Q=1.2$. Hence,

for large samples, Var $\bar{y}$ will be smaller than Var y'. We find in fact that $v(\bar{y}) = 12.7$, where the formula for $v(\bar{y})$ is the usual one (see p. 119 in [1]) except that the finite sampling correction has been omitted here for the purpose of comparison with $v(y') = 113.7$ computed from equation (35). The formula for $v(\bar{y})$ is based on estimating the terms in equation (23), and similarly an estimate of $v(y')$ may be obtained by estimating the terms in equation (24). The estimate of $v(y')$ obtained by this method is $v(y') = 112.0$, which is rather close to the unbiased estimator obtained using equation (35).

In the previous example, we found that y' was better than $\hat{y}$, but the variance of $\bar{y}$ was much smaller than that of y'. We shall now consider a numerical example in which y' is better than both $\hat{y}$ and $\bar{y}$. The production consists of four (x, y) pairs: (2, 2), (2, 6), (4, 6), (6, 10), and a random sample of two pairs is drawn. In this case, $B = 1.45$, $\overline{R} = 1.79$, $Q = 1.71$. Since we are dealing with a sample size of 2, the large sample formulas would not give good approximations. The exact numerical results are presented in the following table, in which we give bias and variance for each of the four estimators of $\overline{Y}$; viz., $\tilde{y}$, $\hat{y}$, y', and $\bar{y}$.

Number of Strata K / Estimate	Mean	Bias	Variance	Mean Square Error					
	1	1	1	1	5	10	20	50	100
$\tilde{y}$	6.09	.09	.909	.917	.189	.098	.053	.026	.017
$\hat{y}$	6.27	.27	2.233	2.407	.520	.297	.185	.118	.096
y'	6.00	.00	.563	.563	.112	.056	.028	.011	.006
$\bar{y}$	6.00	.00	2.667	2.667	.533	.267	.133	.053	.027

We see that y' has a smaller variance and bias than the other estimators. Also, if we assume that there are a number of strata (k) that have the same (x, y) population values, then the relative advantage of using y' for each stratum, rather than the other estimates, increases as k, the number of strata, increases.

7. SOME SPECIAL CASES

In Section 4.2, we saw that the large sample variance for y' is n^{-1} Var $(y - \overline{R}x)$, and the large sample variance of $\bar{y}$ is n^{-1} Var $(y - Qx)$. The estimate $\hat{y}$ is consistent only in the case where $\overline{R} = Q$, and the variance of $\hat{y}$ is n^{-1} Var $\{(y - \overline{R}x)(\overline{X}/x)\}$. The comparison of $\hat{y}$ with y' and $\bar{y}$ (when $\overline{R} = Q$) will depend on the form of the conditional variation $E\{(y - \overline{R}x)^2 | x\}$ of y around $\overline{R}x$, given x, and the distribution of x. We shall study several forms of $E\{(y - \overline{R}x)^2 | x\}$ or of the conditional variation $E\{(r - \overline{R})^2 | x\} = g(x) = E\{[(y - \overline{R}x)/x]^2 | x\}$ of r given x.

We shall also make some assumptions that are often quite reasonable in many situations where ratio estimators are appropriate. It has been pointed out (see e.g., p. 196 in [4]) that there will be no bias, or only a trivial bias, in ratio estimators "when the regression of x and y is reasonably well represented by a straight line through the origin." If the line of relationship between y and x departs in a significant way from going through the origin, the variance of

the ratio estimators may be large. Because of this possibility of large variance and/or bias, the use of ratio estimators seems most appropriate when the line of relationship does not depart in a significant way from going through the origin. We shall therefore assume in this section that the expected value of y given x, $E(y|x)$, may be described by a line going through the origin. In other words, we shall restrict our study in this section to the special case where $E(y|x)=ax$, where a is an unknown positive constant. In this case, the three ratio estimators y', $\bar{y}$, and $\hat{y}$ are unbiased ([1], p. 123), and the exact formulas for the variance and efficiency of these three estimators may be obtained.

In the case where the form of the conditional variation $E\{(r-a)^2|x\}=g(x)$ of r given x is known, the most efficient linear unbiased estimate of $\overline{Y}$ is

$$y'' = \overline{X} \sum_{i=1}^{n} [r_i/g(x_i)] \Big/ \sum_{i=1}^{n} [1/g(x_i)].$$

The exact variance of y'' is

$$\text{Var}\,(y'') = \overline{X}^2 E\left\{1 \Big/ \sum_{i=1}^{n} [1/g(x_i)]\right\} \text{ (we assume } g(x_i) > 0). \tag{38}$$

Since y'' has a smaller variance than any other linear unbiased estimate (i.e., linear function of r_i for which the conditional expectation of the function, given $x_1, x_2, \cdots, x_n$, is equal to $\overline{Y}$), we may write the formulas for the efficiency of the ratio estimators relative to y''. When n approaches infinity, we have that

$$n\ \text{Var}\,(y'') \approx \overline{X}^2/E\{1/g(x)\}, \tag{39}$$

where $E\{1/g(x)\}$ is the expected value of $1/g(x)$ when a single x value is drawn at random from the population.

We shall now determine the exact variance of the ratio estimators. Writing

$$\bar{y} = \overline{X} \sum_{i=1}^{n} x_i r_i \Big/ \sum_{i=1}^{n} x_i,$$

we have that $E(\bar{y}|x_i)=\overline{Y}$,

$$\text{Var}\,(\bar{y}\,|\,x_i) = \overline{X}^2 \sum_{i=1}^{n} g(x_i) x_i^2 \Big/ \left(\sum_{i=1}^{n} x_i\right)^2, \tag{40}$$

and

$$\text{Var}\,(\bar{y}) = \overline{X}^2 E\left[\sum_{i=1}^{n} g(x_i) x_i^2 \Big/ \left(\sum_{i=1}^{n} x_i\right)^2\right], \tag{41}$$

where $E(z|x_i)$ and $\text{Var}\,(z|x_i)$ are the conditional mean and variance, respectively, of z given $x_1, x_2, x_3, \cdots, x_n$, and $\text{Var}\,(z)$ is the unconditional variance of z.

Writing

$$\hat{y} = \overline{X} \sum_{i=1}^{n} r_i/n,$$

we have that $E(\mathcal{Y}|x_i) = \overline{Y}$,

$$\text{Var}(\mathcal{Y}|x_i) = \overline{X}^2 \sum_{i=1}^{n} g(x_i)/n^2, \tag{42}$$

and

$$\text{Var}(\mathcal{Y}) = \overline{X}^2 E[g(x)]/n. \tag{43}$$

Writing

$$y' = \overline{X}\bar{r} + n\left(\sum_{i=1}^{n}(r_i x_i/n) - \bar{r}\bar{x}\right)\Big/(n-1)$$

$$= \sum_{i=1}^{n} r_i\{(\overline{X}/n) + [(x_i - \bar{x})/(n-1)]\},$$

we have that

$$\text{Var}(y'|x_i) = \sum_{i=1}^{n} g(x_i)\{(\overline{X}/n) + [(x_i - \bar{x})/(n-1)]\}^2, \tag{44}$$

and

$$\text{Var}(y') = \overline{X}^2 E[g(x)]/n + E\left[\sum_{i=1}^{n} g(x_i)(x_i - \bar{x})^2\right]\Big/(n-1)^2$$
$$+ 2\overline{X}E\left[\sum_{i=1}^{n} g(x_i)(x_i - \bar{x})\right]\Big/n(n-1). \tag{45}$$

Writing the sample mean

$$\bar{y} = \sum_{i=1}^{n} r_i x_i/n,$$

we have that

$$E[(\bar{y} - \overline{Y})^2|x_i] = E\{[(\bar{y} - a\bar{x}) + (a\bar{x} - \overline{Y})]^2|x_i\}$$
$$= \sum_{i=1}^{n} g(x_i)x_i^2/n^2 + a^2(\bar{x} - \overline{X})^2, \tag{46}$$

and

$$\text{Var}(\bar{y}) = \{E[g(x)x^2] + a^2\,\text{Var}(x)\}/n, \tag{47}$$

where $E[z|x_i]$ is the conditional expectation of z given $x_1, x_2, x_3, \cdots, x_n$, and $E[g(x)x^2]$ is the expected value of $g(x)x^2$ when a single x value is drawn at random from the population.

When n approaches infinity, we have the following results:

$$\begin{aligned} n \text{ Var } (\bar{y}) &\approx n \text{ Var } (y') \approx E[g(x)x^2] \\ Eff(\bar{y}) &\approx Eff(y') \approx \bar{X}^2/E[1/g(x)]E[g(x)x^2] \\ Eff(\hat{y}) &\approx 1/E[1/g(x)]E[g(x)] \\ Eff(\bar{y}) &\approx \bar{X}^2/E[1/g(x)]\{E[g(x)x^2] + a^2 \text{ Var } (x)\}, \end{aligned} \tag{48}$$

where $Eff(z)$ denotes the efficiency of z relative to y''. Hence when n is large, the efficiency of $\bar{y}$ and y' are approximately equal.

Let us now consider some of the special forms of the function $g(x)$ mentioned in the previous section; viz., (1) $g(x)=b/x^2$, (2) $g(x)=b/x$, and (3) $g(x)=b$.

In case (1),

$$\begin{aligned} Eff(\bar{y}) &\approx Eff(y') \approx \bar{X}^2/E[x^2] \\ Eff(\hat{y}) &\approx 1/E[x^2]E[1/x^2]. \end{aligned} \tag{49}$$

Hence, $Eff(\bar{y}) \approx Eff(y') \approx 1/(1+D^2)$, where D is the coefficient of variation of x; that is $D=\sqrt{[E(x^2)-\bar{X}^2]/\bar{X}^2}$. We therefore see that $\bar{y}$ and y' are almost efficient when the coefficient of variation D of x is small, but that these estimators may be inefficient when D is large. For example, when $D=1.01$ the large sample efficiency of $\bar{y}$ and y' is less than 50%. The large sample efficiency of $\bar{y}$ and y' will be greater than the efficiency of $\hat{y}$ if and only if $\bar{X}^2E[1/x^2] \geq 1$ in case (1).

In case (2),

$$\begin{aligned} Eff(\bar{y}) &\approx Eff(y') \approx 1 \\ Eff(\hat{y}) &\approx 1/E[x]E[1/x]. \end{aligned} \tag{50}$$

Hence, the estimators $\bar{y}$ and y' are efficient in case (2).

In case (3),

$$\begin{aligned} Eff(\bar{y}) &\approx Eff(y') \approx \bar{X}^2/E[x^2] = 1/(1+D^2), \\ Eff(\hat{y}) &\approx 1. \end{aligned} \tag{51}$$

Hence, the estimator $\hat{y}$ is efficient. The efficiency of $\bar{y}$ and y' will depend on the size of the coefficient of variation D of x in case (3) as well as in case (1).

The large sample efficiency of $\bar{y}$ and y' will be equal to one if and only if $X^2=E[1/g(x)]E[g(x)x^2]$. In the case where Var (x) and Var (r/x) are greater than zero, this equality holds if and only if $g(x)=b/x$ (case (2)). In other words, $\bar{y}$ and y' are efficient only in case (2). This result follows from an application of Schwartz's inequality.

The large sample efficiency of $\hat{y}$ will be equal to one if and only if $E[1/g(x)]E[g(x)]=1$. In the case where Var (x) and Var $(r|x)$ are greater than zero, this equality holds if and only if $g(x)=b$ (case (3)). In other words, $\hat{y}$ is efficient only in case (3).

If the form of the function $g(x)$ is not known, it may be possible to determine its approximate form by an examination of the scatter diagram. With

this information, it will be possible to approximate the most efficient estimator y''. For example, if an examination of the sample scatter diagram indicates that $E(y|x)=ax$ and $g(x)=b/x$ (case (2)), then the estimator $\bar{y}$ or y' might be best.

8. ON SOME REGRESSION ESTIMATORS

In the preceding sections, we have examined the precision of three ratio estimators. A method for obtaining a large variety of unbiased ratio and regression estimates has been presented by M. R. Mickey [7]. The unbiased regression estimators might be more useful when the line of relationship between x and y departs in a significant way from going through the origin. It might be worthwhile to study the efficiency of some of the ratio and regression estimators obtained in [7].

In this section, we shall study only estimators of the form $\bar{y}-(\bar{x}-\overline{X})z$, where z is a function of the sample pairs (x_i, y_i) which converges stochastically to some finite value, say Z. For example if $z=\bar{r}$, then the estimate reduces to the estimator y' (when n is large), and $Z=\overline{R}$. If $z=\bar{y}/\bar{x}$, then $\bar{y}-(\bar{x}-\overline{X})(\bar{y}/\bar{x})=\bar{y}\overline{X}/\bar{x}=\bar{y}$, and the value of Z is $Q=\overline{Y}/\overline{X}$.

Since z is assumed to converge stochastically to a finite limit, estimators of this general form will estimate $\overline{Y}$ consistently. The limiting distribution of

$$\sqrt{n}[\bar{y}-(\bar{x}-\overline{X})z-\overline{Y}] = \sqrt{n}[\bar{y}-(\bar{x}-\overline{X})Z-\overline{Y}-(z-Z)(\bar{x}-\overline{X})]$$

will be the same as the distribution of $\sqrt{n}[\bar{y}-(\bar{x}-\overline{X})Z-\overline{Y}]$, since $\sqrt{n}(z-Z)(\bar{x}-\overline{X})$ converges stochastically to zero. Hence, when n is large, the variance of $\bar{y}-(\bar{x}-\overline{X})z$ is approximately

$$n^{-1}(\sigma_y^2+Z^2\sigma_x^2-2Z\ \mathrm{Cov}_{yx}). \tag{52}$$

This variance is minimized when $Z=\mathrm{Cov}_{yx}/\sigma_x^2=\rho\sigma_y^2/\sigma_x^2$, where $\mathrm{Cov}_{yx}=C_{yx}$ and $\rho=C_{yx}/\sigma_x\sigma_y$.
If

$$z=\sum_{i=1}^{n} y_i[x_i-\bar{x}]\Big/\sum_{i=1}^{n} x_i[x_i-\bar{x}],$$

then the asymptotic variance of the estimator $y'''=\bar{y}-(\bar{x}-\overline{X})z$ is minimized. This minimum asymptotic variance is equal to $[\sigma_y^2-C_{yx}^2/\sigma_x^2]/n$. Hence the large sample efficiency of the estimator y' (relative to y''') is

$$[\sigma_y^2-C_{yx}^2/\sigma_x^2]/[\sigma_y^2+\overline{R}^2\sigma_x^2-2\overline{R}C_{yx}]=(1-\rho^2)/(1+F^2-2F\rho), \tag{53}$$

and the efficiency of $\bar{y}$ is

$$[\sigma_y^2-C_{yx}^2/\sigma_x^2]/[\sigma_y^2+Q^2\sigma_x^2-2QC_{yx}]=(1-\rho^2)/(1+H^2-2H\rho), \tag{54}$$

where $H=Q\sigma_x/\sigma_y$ and $F=\overline{R}\sigma_x/\sigma_y$ (see p. 148 in [1]).

The reader will note that the most efficient estimator, among those considered in this section, was the usual regression estimator. Although we did not assume in this section that the type of population might be described by a regression model (i.e., we did not restrict the discussion to the case where

$E[y|x]=ax+b$), we find that the standard regression estimator has a smaller asymptotic variance than other estimators of the form we have considered. It can also be seen that there are many unbiased regression estimators that are relatively efficient; i.e., unbiased and equivalent to y''' when n is large.

If we do assume that the regression model holds, some further results may be obtained. When $E[y|x]=ax+b$, then

$$Q = \bar{Y}/\bar{X} = a + b/\bar{X},\ \bar{R} = E[y/x] = a + bE[1/x],\ E[y(x - \bar{X})] = aS_x^2.$$

The estimator $\bar{y}$ will be more precise than y' if and only if

$$|\bar{R} - C_{yx}/\sigma_x^2| > |Q - C_{yx}/\sigma_x^2| \quad \text{(see Section 3.1)}; \tag{55}$$

i.e., if and only if

$$\begin{aligned} |bE[1/x]| &> b|\bar{X}|, \\ \bar{X}E[1/x] &> 1 \qquad (b \neq 0). \end{aligned} \tag{56}$$

Since this is always the case when $b\neq 0$ and $\sigma_x^2>0$, we see that $\bar{y}$ will always be more precise than y' when $b\neq 0$. In the preceding section, we considered the special case when $b=0$, and saw that the estimators $\bar{y}$ and y' had the same precision in that case.

We have seen that when $E[y|x]=ax+b$, the estimator y' is never more precise than $\bar{y}$. This fact may also be obtained using the following somewhat more general consideration which may help to throw some more light on the relationship between the estimators. In general, we have that

$$Q - \bar{R} = \bar{Y}/\bar{X} - E[r] = \{\bar{Y} - \bar{X}E[r]\}/\bar{X} = C_{rx}/\bar{X},$$

where C_{rx} is the covariance between r and x. When $E[y|x]=ax+b$, then

$$C_{rx} = bE[(x - \bar{X})/x] = b\{1 - \bar{X}E[1/x]\}.$$

Since $\bar{X}E[1/x]\geq 1$, we see that $C_{rx}\geq 0$ if and only if $b\leq 0$. Hence, $Q\geq \bar{R}$ if and only if $b\leq 0$. Since $b=\bar{Y}-a\bar{X}=(Q-a)\bar{X}$, we find that when $b\leq 0$, then $Q\leq a$, and $a\geq Q\geq \bar{R}$. When $b\geq 0$, then $a\leq Q$, and $a\leq Q\leq \bar{R}$. Since $C_{yx}/\sigma_x^2=a$ in the regression model, we see that the value of $\bar{R}$ is never nearer to C_{yx}/σ_x^2 than is the value of Q. In fact, we see that when Q is too large (i.e., larger than the optimum value C_{yx}/σ_x^2), then $\bar{R}$ will be even larger than Q; and if Q is too small then $\bar{R}$ will be even smaller. Hence when n is large, the variance of y' is never nearer to the minimum variance (the variance of y''') than is the variance of $\bar{y}$.

REFERENCES

[1] Cochran, William G., *Sampling Techniques*, New York: John Wiley and Sons, Inc., 1953.

[2] Cook, M. B., "Bi-variate K-statistics and cumulants of their joint sampling distributions," *Biometrika*, 38 (1953), 179–95.

[3] Fisher, R. A., "Moments and product moments of sampling distributions," *Proceedings of the London Mathematical Society*, 30 (1929), 199–238.

[4] Hansen, Morris H., Hurwitz, William N., and Madow, William, *Sample Survey Methods and Theory*, Vol. 1, New York: John Wiley and Sons, Inc., 1953.

[5] Hartley, H. O., and Ross, A., "Unbiased ratio estimators," *Nature*, 174 (1954), 270–1.
[6] Kendall, Maurice G., *The Advanced Theory of Statistics*, Vol. 1, London: Charles Griffin and Co., Ltd., 1947.
[7] Mickey, M. R., "Some finite population unbiased ratio and regression estimators," unpublished report.
[8] Raj, Des, "Ratio estimation in sampling with equal and unequal probabilities," *Journal of the Indian Society of Agricultural Statistics*, 6 (1954), 127–39.
[9] Robson, D. S., "Applications of multivariate polykays to the theory of unbiased ratio-type estimation," *Journal of the American Statistical Association*, 52 (1957) 511–22.

ON NONCOVERAGE OF SAMPLE DWELLINGS*

LESLIE KISH AND IRENE HESS
Survey Research Center, University of Michigan

We present an investigation of the extent of noncoverage that can occur in area sample surveys of dwelling units in the United States, and of some relevant factors associated with noncoverage. We describe changes in field procedures which resulted in drastic improvements in dwelling unit coverage. Our noncoverage, formerly somewhere in the neighborhood of 10 per cent, was roughly the same magnitude as the total of all nonresponses; therefore, its reduction was important. In one year we brought our coverage rate to about 97 per cent of the number of occupied dwelling units reported by the Census Bureau. Thus noncoverage, as a source of errors of nonobservation in our surveys, has been greatly reduced in magnitude and importance. The potential bias is shown to be a function of the type of estimation used: large for simple expansions, less for means and usually even less for comparisons of subclass means. The amount of noncoverage is a function of the essential operating conditions used by a specific organization; however, the usefulness of the methods and results—and the nature of the problem—are quite general to surveys.

TYPES OF NONSAMPLING ERRORS IN SURVEYS

To place the errors of noncoverage in proper perspective we present below a brief classification of the types of nonsampling errors that arise in sample interview surveys from households in the United States.

1. *Errors of observation* consist of the errors due to obtaining and recording an observation incorrectly. These may be: *response errors* or *processing errors.*
2. *Errors of nonobservation* are due to failures to obtain observations on some segment of the survey population: a) *Errors of nonresponse* include failures to obtain an observation (measurement) on an observational unit designated for the sample. These may be classified as refusals, not-at-homes, incapacitated, lost schedules, etc. Nonresponse may be total or partial. b) *Errors of noncoverage* are due to failures to include in the sampling frame some members or some segments of the population as defined for the survey. These include various errors made in the selection process, in office routine, in field instructions, and in field procedures. The reverse of these are errors of over-coverage: including in the sample elements which do not belong to the sampling unit or do not belong to the population.

Failure to include sampling units in the sampling frame results in a zero probability of selection and in the exclusion of these units from the survey results. However, in this subclass of noncoverage we do *not* refer to any segments of some larger population that have been *deliberately* and *explicitly* excluded from the survey population. This kind of exclusion arises either from considerations

* Some of this was presented at the Annual Meeting of the American Statistical Association in Atlantic City on September 13, 1957.

of the survey objectives or because of practical difficulties. For example, in surveys of attitudes of the United States population, persons under 21 years of age are often excluded explicitly and deliberately; residents of institutions of various kinds are often excluded because of practical difficulties. This kind of explicit exclusion differs both in intent and in result from noncoverage caused by failures in the procedures. The former are part of the stated explicit "*essen-sential conditions of the survey.*"[1] If they are identifiable it is sometimes possible to estimate their effects.

We separate noncoverage from nonresponse because of the great practical differences in the methods for and difficulties in measurement. The extent of nonresponse can usually be estimated well and simply from checks of the survey returns against the sampling units designated in the selection procedures; and it is common these days to find the percentage of nonresponse attached to results of probability samples.

However, noncoverage is difficult to estimate in many survey situations, and coverage estimates for area samples of dwelling units are scarce. The extent of noncoverage in these situations can be estimated only against some check obtained *outside* the survey procedure itself. This check can come either from an auxiliary investigation attached to the survey, or from data obtained from independent sources. In either case, the task of checking is an arduous task, seldom performed.[2]

From the scarcity of literature on noncoverage, it would be wrong to conclude that it is not an important source of survey errors. It is true that there are many sampling situations where the existence of a clear, convenient, and (nearly) complete frame will allow the sampler to keep noncoverage low or to estimate it easily and precisely. For example, it may be easy to sample the employees of a firm, the telephone subscribers of a city, or the students of a school. The national registers of some European countries provide easy and relatively complete frames of their populations. In these situations the estimation of noncoverage may also be relatively simple.

This article investigates a difficult and important practical problem: obtaining coverage of the dwelling units of the United States by means of an area sample. The task involves some form of listing of the dwellings by the interviewer. Practitioners agree that this job is a difficult one to do well—that is, to a close approximation of complete coverage. Even the United States Census of 1950 missed 2 or 3 per cent of the dwellings (see reference in footnote 2) in spite of its excellent organization and its concentration on complete coverage.

PROBLEMS OF ESTIMATING THE NONCOVERAGE

The estimation of errors of noncoverage generally entails one of two difficult alternatives. The first calls for a quality check by procedures which are suffi-

[1] Terms such as "essential conditions of a survey" or "true value" have been defined with some (but incomplete) uniformity in several fine texts on survey sampling. See the volumes by Cochran, Deming, Sukhatme and Yates; also Morris H. Hansen, William N. Hurwitz, William G. Madow *Sample Survey Methods and Theories*, John Wiley and Sons, New York, 1953. Vol. II, Chap. 12.

[2] Both kinds of checks can be illustrated by investigations on the noncoverage of the 1950 Census of the United States population. The remarkable auxiliary project of the Census Bureau is summarized by A. Ross Eckler in "Extent and character of errors in the 1950 census," *The American Statistician*, December, 1953. An investigation against independent sources is: "The population of the United States in 1950 classified by age, sex, and color—a revision of census figures," Ansley J. Coale, *Journal of the American Statistical Association*, March, 1955, pp. 16–54.

ciently better to provide the "true value" against which the survey results can be compared. The quality check is necessarily expensive, both because the separate and better procedure is generally costly, and because the check has to be on a large enough scale to measure errors which are generally rare and "spotty." The alternative procedure calls for a reliable estimate from an *outside* source. One difficult condition is that the outside estimate of the total must be based on the same units not only by formal definition, but also operationally; the differences in the two definitions must be small enough to allow for meaningful comparisons between the two sources.

The Census Bureau provides annual estimates of the number of households, equivalent to the number of occupied dwelling units, in the United States.[3] It is against the Census Bureau's national estimates that comparisons in this paper are made. Our technique consists basically in comparing with the Census national estimate the occupied dwellings in the Survey Research Center (SRC) samples multiplied by the inverse of the over-all sampling fraction (probability).[4] Hence, the amount of noncoverage in the Census estimates is additive to our noncoverage estimates. These, moreover, are for the *net* noncoverage hiding whatever overcoverage may exist in them.

In the light of the operating conditions, it seems to us that occupied dwelling units provide the most practical basis for comparisons of our surveys with current data from the Census Bureau. Comparisons based on persons seem somewhat impractical both because these are not collected uniformly in our surveys, which deal with many different populations, and because of the difference in the bases of the two sources. In particular, the population of total persons in the United States is larger than the population living in dwelling units to which our samples are confined—larger by an appreciable number, which is difficult to estimate.

While occupied dwelling units provide a useful basis for our comparisons, that basis is not without faults and doubts—because of possible differences in actual operations despite the formal similarity of definitions. Although estimates of the over-all coverage are useful, they are not entirely precise; they are subject both to some bias of measurement and to sampling errors. The raw figures of our investigation for the years 1950–1955 yield estimates of noncoverage somewhere between 11 and 13 per cent. However, after some adjustments, our best guess is that noncoverage during the period of our "old methods" of block listings was about 10 per cent. With present procedures we estimate that noncoverage for the past two years has been about 3 per cent. Furthermore, with new controls and with wider use of changed field procedures we may be able to reduce this figure still further.

The improvement during the last two years was beyond our fondest expectations. For survey practitioners it is a rare experience to reduce so drastically an important source of errors, and to find new procedures which are so

[3] U.S. Bureau of the Census, *Current Population Reports:* Series P-20, Nos. 68, 75, 76, "Households and Families by Type."

[4] We wonder if now is the time for: a) A U.N. Statistical Sub-Commission report to standardise terms and recommend usage in the area of non sampling errors similar to the excellent and effective "The Preparation of Sampling Survey Reports," U.N. Statistical Paper—Series C, No. 1; and b) reports of probability samples to include estimates of noncoverage along with nonresponse.

much more effective without considerably increasing costs. Moreover, the sudden change came after years of attempts at improvement with changes of instructions—attempts which looked good on paper but failed in the field. These considerations prompt us to share our experiences in the expectation that some of the problems are common to survey organizations in spite of differences among them.

Beyond an over-all estimate of noncoverage we undertook a painstaking investigation of its major components in different parts of the population. The large differences in noncoverage provide clues to some of the possible effects of noncoverage on survey estimates, especially for allocation of efforts, in the office and in the field, for the elimination of the major sources of noncoverage. The analysis consists of two parts, which are distinguished by the nature of the unit (block or place), used in the analysis.

BLOCK ANALYSIS

The block analysis consists of 629 blocks in 50 cities of over 50,000 population, for which block data were available from the 1950 Census ("block cities"). The basic estimate of coverage consisted of the ratio for each sample block:

$$\frac{\text{Addresses listed by SRC interviewers} \times \text{Dwelling units per sample address}}{\text{Dwelling units in the 1950 Census} \times \text{Growth factor for the city}}$$

The addresses listed within the sample blocks needed two kinds of corrections. First, although the interviewer is asked to list dwelling units, he cannot do this perfectly by observing the buildings from the outside. He actually enters on the listing sheet visible listing units which we call "sample addresses." Some additions and deletions take place during interviewing when the interviewer obtains more information at the sample address. The net correction factor of dwelling units per sample address averaged about 1.04, resulting from four classes of discrepancies: additional dwelling units at a sample address, sample addresses missed during the listing operation and added by means of the "half-open interval" technique during interviewing, newly constructed dwellings, and listed sample addresses which were found at the time of interviewing to contain no dwellings. The first three classes were added and the last class subtracted to translate sample addresses into dwelling units. These correction factors were applied in several groups of cities classified by size and region.

Secondly, the growth factor was used to adjust for the average growth of the city between 1950 and the date of the listing—mostly within a year or two, and not later than 1954. The growth factor took into consideration the rate of growth of the city between 1940 and 1950 (based on Census data) and the rate of growth since 1950 as estimated from building permit data.

Unfortunately these correction factors are far from precise. Hence, the overall coverage figure is listed as 96 per cent in Table 513, while it should be nearer 90 per cent. This lack of precision also may affect somewhat the comparisons between the various subclasses in Table 513; nevertheless the larger differences found there are real, we believe.[5]

[5] a) Most of our data originate within a year or two of the Census of 1950; hence, the uncertainties due to changes are rather confined. b) We computed a standard error of about $s = 0.4$ per block for these ratios. The simple formula of $s/\sqrt{n}$, where n is the number of blocks for an estimate based on about 70 blocks points to a standard error of 0.05, and for 200 blocks a standard error of about 0.03.

The sample blocks were selected with probabilities proportionate to Census size, and the over-all sampling fraction was determined so as to yield a self-weighting sample of dwelling units for the United States. Hence, the analysis of these blocks reflects the condition in the population of "block cities."

Some of our results are presented in Table 513; the 629 blocks were classified in four different ways denoted, respectively, by Sections A, B, C and D.

The coverage is higher in the blocks with higher economic ratings. The differences are not small: roughly six per cent higher in the "medium" and 12 per cent higher in the "high" group than in the "low" group. The coverage data are given first for the blocks classified (A) on the basis of mean rental or property value, as reported in the Census Block Statistics bulletins. The second classifica-

TABLE 513

COVERAGE RATES* BY CHARACTERISTICS OF THE BLOCK
(Numbers in parentheses give the number of blocks on which the rate is based)

Block Characteristics	City Size (in Thousands)			
	Total	50–99	100–249	250 and over
All Blocks	0.96 (629)	0.92 (154)	0.94 (114)	0.99 (361)
A) *Census Average of Property Value or Rental*				
High	1.07 (92)	1.04 (14)	0.94 (21)	1.13 (57)
Medium	1.00 (308)	0.96 (82)	0.92 (48)	1.03 (178)
Low	0.93 (229)	0.88 (58)	0.94 (45)	0.94 (126)
B) *Average of Economic Ratings Assigned by SRC Lister*				
High	1.05 (117)	1.18 (22)	1.05 (26)	1.00 (69)
Medium	0.96 (283)	0.90 (56)	0.91 (48)	0.99 (179)
Low	0.94 (229)	0.89 (76)	0.92 (40)	0.99 (113)
C) *Prevalent Type of Structures in Block*				
Single dwellings	0.93 (312)	0.88 (104)	0.90 (70)	1.00 (138)
2 to 7 dwellings	0.95 (46)	0.94 (24)		0.97 (22)
Apts. above stores	0.94 (56)	0.91 (16)		0.96 (40)
Apartment houses	1.04 (70)	—	—	—
Mixtures of several classes	0.99 (145)	1.06 (24)	0.94 (25)	1.00 (96)
D) *Nature of Block Boundaries*				
Simple rectangle or polygon	0.95 (400)	0.92 (84)	0.91 (65)	0.98 (251)
Composite block (combination of two or more blocks)	0.96 (157)	0.94 (42)	0.98 (36)	0.96 (79)
Irregular shape and/or large area	1.00 (72)	0.89 (28)	0.94 (13)	1.16 (31)

* $\frac{\text{Addresses listed by SRC interviewers} \times \text{Dwelling units per sample address}}{\text{Dwelling units in the 1950 Census} \times \text{Growth rate for city}}$

(Note: Coverage ratios are subject to an estimation bias and to sampling error. The figures are to be used solely for comparisons and not as actual measures of coverage.)

tion (B) is based on examination of the economic ratings assigned by the lister to the individual dwellings.[6]

The coverage was better in the largest cities than in those of moderate size. Poor coverage comes from the moderate-sized cities, and there from the "medium" and especially from the "low" blocks.

[6] To appraise these effects on the entire sample, it must be pointed out that the size of each stratum is not proportional to the number of blocks: the rates of selection of the high, medium, and low economic blocks were in the proportion of 4:2:1 in the earlier period and 2:2:1 later. The strata as they stand represent about 10 per cent, 40 per cent and 50 per cent, respectively, of all dwelling units in cities of 50,000 and over.

The blocks were also classified in the office according to the prevailing type of structures in the block (C). As expected, coverage was poor in blocks where small multi-dwelling structures of two to seven dwelling units predominate, or where many apartments are located above stores. It seems that coverage is very good in blocks containing chiefly apartment houses—probably because of the systematic design of these larger buildings. On the other hand, we were surprised by the poor coverage in blocks where the listing showed chiefly single dwellings; perhaps these blocks contain many scattered back entrances and second floor apartments which remain unnoticed by the lister.

The nature of the block boundaries (D) was of interest to us. Some people supposed that much of our trouble came from blocks that were large and irregularly shaped, where the interviewers had special difficulties in finding all of the dwellings. We found no evidence for that supposition. The "composite" block consists of two or more adjacent blocks grouped to provide a block of "sufficient" measure of size; these blocks also failed to reveal any unusually large noncoverage rates. It seems that the majority of missed dwellings come from ordinary-sized, well-defined regular blocks, in moderate-sized cities, which appear to the lister to be filled with ordinary single-dwelling homes.

ANALYSIS BY PLACE

The analysis of places (cities, towns and villages) was based on two types of data. In the "block cities" the coverage ratios of the individual blocks (described above) were averaged. For the smaller places, where the sample blocks were selected systematically from a numbered map of all blocks, the coverage ratio was computed as:

$$\frac{\text{DU's listed in sample blocks}}{\text{DU's in 1950 Census}} \times \frac{\text{Total number of blocks in town}}{\text{Number of sample blocks in town}}$$

The coverage ratio for each place was corrected for the factors of listing (dwellings per sample address) and of growth. The 113 places (cities and towns) which enter into the analysis are not our entire sample of places for that period, but we believe the conclusions are not thereby vitiated. They are all those for which meaningful comparison could be made; for other places either some data were lacking or the differences in city boundaries made comparison hazardous.

In Table 515 we may observe appreciable differences in coverage of the different size classes. The coverage figures for the eight size classes are far from precise; nevertheless, they suggest that medium-sized cities, those from 10,000 to 250,000 population, had the poorest coverage. For this conclusion there is supporting evidence from checks by the 1950 Census Post Enumeration Survey enumerators of SRC listings. The percentage coverage showed similar variation in population classes, with lowest coverage for the medium-sized cities. Those coverage rates were: towns under 2,500, 99 per cent; cities 2,500 to 50,000, 94 per cent; cities 50,000 to 500,000, 88 per cent; cities 500,000 and over, 96 per cent.[7] (See footnote 20.)

[7] In addition to the average coverage rates we included in Table 515 another measure for comparing the coverages in different size classes of cities: the proportion of cities in each size class where the coverage rates fell below 1.00. This is seen to be highest in the medium-sized cities.

The standard deviations within size classes is about 0.12 per city (on the average) for cities over 50,000 and

TABLE 515

COVERAGE RATES BY SIZE OF PLACE (CITY OR TOWN)

Item	1950 Population (in thousands)							
	1.0–2.4	2.5–9.9	10–24	25–49	50–99	100–249	250–499	500 & over
Coverage Rates	1.03	1.10	0.94	0.89	0.94	0.92	1.05	1.00
Standard Error	0.06	0.07	0.08	0.15	0.04	0.02	0.06	0.03
Number of Places	10	32	13	8	16	11	6	17
Proportion of Places with coverage less than 1.00	0.40	0.31	0.62	0.75	0.79	0.82	0.33	0.53

Note: Coverage rates, etc.—same as under Table 513.

The poor performance in the medium-sized cities came as an unexpected result—which we later put to good use. *Post hoc*, we reason that the cities of medium size are those in which irregular crowding into small multi-unit structures (two to seven dwellings) is most prevalent; and this type of structure is the hardest to list well (see Table 513). This situation is in contrast with more widely spaced, one-family homes in small towns, and with the well regulated character of apartment houses and of large real estate developments in the larger cities.

Is the poor coverage in medium-sized cities due chiefly to differences in growth rate? This supposition was contradicted by further analysis. But there is an interesting relationship between growth rates and coverage (Table 516)! The coverage is poor in both the slowest and fastest growing cities, with best coverage in cities with moderate growth of 0.01 to 0.02 annually. We investigated this problem only in cities of 50,000 and over; in the others the data on growth seemed much too unreliable for testing. Perhaps the slowly growing old cities have a larger proportion of low-income areas, and in the very fast growing cities obstacles are more often found in the form of conversions of old, large houses into multiple dwellings.

There were two results which tend to show that coverage can be influenced strongly by variations in the nature of training and supervision. First, there seemed to be about 5 per cent coverage variation in the listings of different years. Unfortunately, any possible differences of emphasis in training for listing were unplanned, indefinite, and forgotten; hence they could not be correlated with the results. Secondly, a surprising difference was found between the two sets of primary sampling units (psu's) into which our national sample of 66 psu's has been divided. The "A" half-sample was used considerably more frequently, during the years in question, than the "B" half-sample. Hence, the interviewers in the B half-sample had less experience. The results were surprising: the B half-sample had a considerably better coverage, by over 5 per cent.

about 0.34 for the smaller places. The computed standard error is presented for each group. The smaller variation for the larger cities is mostly due to the selection of blocks with probability proportional to Census size which reduces considerably the variation in sample size per block.

TABLE 516

COVERAGE RATES BY AVERAGE GROWTH RATE OF CITY
Cities of 50,000 or more population in 1950

Item	Average Annual Growth Rate				
	Under 0.010	0.010–.014	0.015–0.019	0.020–0.029	0.030 & Over
Coverage Rate	0.88	0.93	0.94	0.91	0.85
Number of places	12	10	8	11	7

Apparently the interviewer is at first careful about the listing instructions he just learned, but becomes careless later.[8]

Examination of the coverage ratios for individual cities disclosed considerable variation from place to place; about half of the noncoverage was attributable to a relatively few places. Some places are more difficult to list than others. Nevertheless, we suspect that a small proportion of the field workers commit a large proportion of the errors, while the majority do a much better job.[9]

LIMITATIONS OF THESE FINDINGS

These results deal chiefly with the problems of noncoverage of a procedure using "block listings" of dwellings in sample blocks with subsequent subsampling. All this was done by interviewers in a widespread national sample. This is an important practical sampling problem; but how relevant are these results to other sampling situations? How can our experiences be compared with those of other organizations? Any projection of our experiences to another situation should give consideration to the following:

A. *Other frames* will result in different coverage. From a clear, convenient, and complete frame it is rather easy to obtain good coverage.

B. For a single *city or county* it is generally easier to obtain good coverage. Immediate and direct supervision and checking of the field work permits easier and tighter control. An example of this is provided by the Detroit Area Study, a sample of dwellings of the city of Detroit plus its surrounding Census-tracted area. The dwellings in about 250 sample blocks were listed by graduate students in the social sciences.[10] In five yearly surveys, 1952–1956, based on three separate listings, the coverage rate was estimated between 95 per cent and 98 per cent.

C. *Entire segments* completely enumerated will lead to good coverage more easily than subsampling of listed segments. This is believed by some experienced samplers, and our experience seems to bear this out.

[8] The Census Bureau also found the old interviewers for the Current Population Survey doing a poorer job, on questions of the labor force, than the new interviewers. See Morris H. Hansen, William N. Hurwitz, Harold Nisselson, and Joseph Steinberg, "The Redesign of the Census Current Population Survey," *Journal of the American Statistical Association*, Sept. 1955, pp. 715–6.

[9] Similarly in the Current Population Surveys: "As can be seen, a relatively small proportion of the interviewers made a large proportion of the errors in coverage." Morris Hansen and Joseph Steinberg, "Control of Error in Surveys," *Biometrics* (1956), XII, pp. 467–468.

[10] R. Freedman, "The Detroit Area Study: training and research laboratory in the community," *American Journal of Sociology*, 59, July 1953, pp. 30–33.

D. A *thorough pre-enumeration* of the entire segment or block before subsampling should produce good coverage. This will involve knocking at most doors and conducting a short pre-interview. This will prove a fruitful procedure in some situations, particularly those where the results of the pre-interview are used for a "double sampling" procedure. However, as a general practice, it has two disadvantages: it is expensive, and some experts believe that it may build up resistance to or bias in the interview.

E. A *half-open interval* procedure was tried for several years. The interviewer was asked to include in the sample not only the sample dwelling but any other dwelling found up to, but not including, the next listed address. The interview schedule for the sample address had two questions dealing first with additional dwellings at the sample address and, secondly, "up to the next listed" address. Similarly fruitless were two attempts to supplement with special searches the listings of selected blocks. These procedures for obtaining good coverage look fine on paper, but our field attempts picked up only a per cent or two of the large noncoverage we knew we had. The failure of our attempts is explained by field experts in terms of the overriding preoccupation of the interviewer with obtaining a good response at the sample address. Perhaps better control of and more emphasis on these procedures could eliminate most of the noncoverage.

F. Could a *different set of written instructions* have increased the coverage substantially? We doubt it. We have worried about noncoverage and rewritten instructions hopefully for eight years only to find that non coverage remained at about the same level, year after year.

G. Could *different training, supervision, and control* produce better coverage with essentially the same listing methods? This question is aimed at more thorough changes in the entire field training and control than the changes in written instructions discussed in E and F. We do not know the answer. Our field force enjoys an enviable reputation. We keep many forms for painstaking control over field procedures which compare favorably with those of most organizations. However, the chief emphasis in the field training and supervision is on the psychological problems of obtaining good cooperation and accurate response from the respondent. Would it be possible to maintain these interviewing skills at their present high level, but at the same time give more attention to the tasks of complete enumeration? Could we have made the same listing methods succeed with more control, continuous quality checks, periodic retraining? We do not know. We decided instead to try some drastic changes in procedures; and these proved so successful that we intend to continue to use them.

CHANGED PROCEDURES

A few words about our organization may lead to a better understanding of our choices of sampling techniques. In order to meet the demand for a constant flow of samples to the field, periodically we draw larger samples which can be

subsampled on relatively short notice for particular studies. We draw five to ten subsamples from each of these larger samples, and in the course of a year or two, we largely exhaust them. The larger samples are redrawn in individual places as they become exhausted. The ratio of the smaller to the larger samples is geared for efficiency in three ways: a) the specialized tasks of sampling are done relatively infrequently and in convenient "slack" periods; b) the entire sample is usually exhausted finally, hence its cost is distributed over several (five to ten) surveys; c) the sample is made to last only a year or two so that it does not become too obsolete. These procedures are adjusted to the nature and size of our organization. Consequently, some of our techniques, which give us substantial benefits, may be inefficient for one-time samples and for other organizations. We use three procedures now:

A. Sampling from city directories is now our preferred procedure in all cities for which an up-to-date directory is available. An area sample in a subsample of cities supplements the directory addresses.
B. A sample of segments, averaging about four dwelling units, is used in rural areas and in some smaller urban areas. We are gradually introducing this procedure in the larger cities also.
C. For some cities, including the largest, directories are not available and we have continued to use our existing prelistings, supplementing these with the half-open interval.[11] We are now changing these places to one of the other two procedures.

The over-all effectiveness of our present procedures, measured[12] by the average coverage rate over several studies conducted recently, was 97 per cent. We consider this strong evidence of significant improvement in our coverage of occupied dwelling units.[13] Because of the consistently high rate over several studies, the change can be attributed to changed procedures rather than merely to improved performance of the interviewers. This inference is justifiable in view of the fact that through eight years of efforts to improve our old procedures—including the use of the half-open interval—our coverage remained at or below 90 per cent. We prepared rough estimates of coverage for our three procedures; these are subject to substantial errors because of lack of data for comparisons. The coverage rates appear to be roughly: 90 per cent with block listings, 95 per cent with city directories, and very close to complete coverage in segments. The 95 per cent coverage in places where we use city directory samples is rather satisfactory in view of the fact that these include chiefly the medium-sized cities where formerly the coverage was particularly low (Table 515), about 85 per cent.

It is one advantage of probability sampling over judgment sampling that its errors can often be measured and sometimes corrected or reduced. Moreover,

[11] In a current study we are replacing the half-open interval technique with an area sample which parallels the city directory block supplement in design and field operation. The results of this experiment are not yet available. Somewhat similar efforts have failed in the past.

[12] This is done by comparing our estimates of occupied dwellings with the Census estimates. See U.S. Census: *Current Population Reports*, Series P-20, No. 68; "Households and families by type," Nos. 68, 75, 76. The noncoverage in Census estimates, 1 to 3 per cent, should be added to ours, for an over-all estimate.

[13] It is difficult to do much better. The Current Population Surveys have outstanding controls and performance; they report a 2 to 4 per cent noncoverage (ref. in footnote 9, p. 473).

its improvements (like other scientific progress) tend to be permanent, cumulative, and communicable. We recognize that our coverage estimates will be subject to sampling fluctuations; that consistent high coverage will depend on continued checks and vigilance; and that our experiences need to be reinterpreted before they are applied to other situations. But we believe that we have established solid bases for continued high coverage rates for our organization.

SAMPLING FROM CITY DIRECTORIES

We are now using city directories as sampling frames in about 60 cities of roughly 3,500 to 1,000,000 population, and they represent about 32 per cent of our national sample. Our procedure[14] has two key features to justify the use of directories for probability samples. First, our samples consist of addresses drawn from the *street guide section of the directories* and not from the alphabetical list of names in the directories. Households are both more mobile and more changeable than dwellings. (One-fifth of the United States households move each year; and the composition of households is subject to many changes as well as to problems of identification.) Secondly, the sample of city directory addresses is supplemented by an area sample of dwelling units missing from the directory, (either missed by the city directory lister or built since the directory listing), so that these too have their proper probabilities of selection.

Neglect of the two key features will result in poor coverage; in a sample selected from householders' names in the Champaign, Illinois, city directory, Ferber found a noncoverage of about 35 per cent, plus an unknown noncoverage due to new construction.[15] However, proper procedures to obtain directory samples of cities have been practiced sporadically for decades. We have used these procedures successfully for several years with some care for the details but without great difficulty. Woolsey[16] describes procedures which, if followed carefully, should enable one to obtain a rather good sample of a single city.

However, to use directories for national samples, more detailed and explicit procedures are needed. When dealing by mail with far-removed interviewers, provision must be made for anticipating and correcting problems and exceptions. We have designed, used and tested such procedures.

Into these general features we introduced two modifications for improving efficiency and quality by taking advantage of our continuous survey operations. First, instead of individual addresses we select clusters of 5, 10, or 20 *lines* from a sample of pages. The sample of lines is converted into a sample of addresses by eliminating blank lines, and by noting the number of different surnames at each residential address. Then this list of addresses is visited by an interviewer to identify nonresidential addresses and addresses outside the survey boundary, to purify and clarify the list, and to assign economic ratings to dwellings (for use in disproportionate sampling). For the most part the list now

[14] Two alternate procedures were tried and found to be impractical. These were: 1) revision of interviewers' listings through a checking operation with the city directories; 2) transcription of city directory listings for entire blocks, then sent to our interviewers for checking and revision. We also considered and rejected proposals to concentrate the address sample and the area supplement in the same blocks.

[15] Robert Ferber, "The Effect of Respondent Ignorance on Survey Results," *Journal of the American Statistical Association*, 51 (December, 1956), p. 578.

[16] Theodore D. Woolsey, "Sampling Methods for a Small Household Survey," Monograph No. 40, (U.S. Public Health Service) (Washington, 1956).

contains well identified dwellings (few corrections remain to be made at the time of interviewing); from it, samples of addresses are selected in clusters of desired sizes. Thus, essentially, our sample is selected in two phases. The first phase sample supplies addresses for five to ten samples drawn, in the second phase, over a period of 10 to 15 months. Thus, our samples are based on up-to-date "listings" prepared generally within a few months of their use.

Secondly, the area supplement is selected in two or three phases. We concentrate the supplement in a stratified subsample of about a fifth of the eligible cities to reduce the necessary effort for what is a relatively small yield. In these cities sample blocks are selected in groups of two or four blocks in larger cities, singly in small towns. The interviewer visits and reports briefly on these blocks and on that basis a subsample of blocks is drawn for each study. For the selected blocks a list of all directory addresses is transcribed. The interviewer looks for omitted addresses and interviews at a designated subsample. We find that the block supplement contributes about 6 per cent of the total directory sample. The directory did not "miss" all these dwellings; some are newly built and some merely represent cases where the directory listed them under different addresses. We found it practical to include these also in the supplement.

Our sampling with directories—with block supplements—is satisfactory, and a great improvement over our previous listing procedures in medium-sized cities, where our listings were poorest. It has two important advantages. First, each address is covered by two operations: listing for the directory, and a check by our field staff; the radical difference between the two procedures reduces the likelihood of noncoverage. Secondly, sampling from directories is generally no more expensive than a good listing—it may be somewhat cheaper.

THE SAMPLE OF SEGMENTS

There is nothing new about the idea of selecting small, compact segments to be completely enumerated on a specified survey.[17] One needs compact segments with good boundaries, but with few dwellings. This is often difficult; in some situations prelisting and subsampling are needed to arrive at a "segment" which is not compact but composed of an average of about four dwelling units selected from a listing of a larger number of dwelling units. We shall describe briefly the more interesting features of our segment procedures.

For more than two years we have used segments averaging about four dwelling units each in the rural areas and in a few urban places in our psu's. Segments are selected in two phases. First we select "chunks" with a measure from about 20 to 30 dwellings, depending upon the particular area. It is easier to achieve definable boundaries for clusters of 30 dwellings than for segments; but sometimes even this requires the help of our interviewers, who provide us with detailed information not available from maps. Each sample chunk is visited in a preliminary field operation to divide it into segments and to supply dwelling unit estimates. Thus our map information is brought up to date and

[17] The arguments for compact segments are made very clearly in W. Edwards Deming, *On the Possible Types of Sampling Unit in the Last Stage of Selection in a Probability Sample*, New York: Advertising Research Foundation, 1955. Among samples using compact segments are the current Population Surveys of the U.S. Census Bureau and the Labor Force Surveys of the Canadian Bureau of Statistics.

we protect ourselves against "surprises" in the form of large-scale new construction (or large trailer camps) at the time of interviewing. This reduces the effect of variations in cluster sizes on a single study. Through subsampling the effects of "variation" are spread over a number of studies.

In the second phase, all dwelling units within the selected segments are listed and located on the sketch of the chunk. In addition to providing the interviewer with a county map on which each chunk is outlined, we prepare a detailed sketch of each chunk showing all segments within the chunk. Duplicate folders are prepared with a sketch mounted in each. One folder is kept by the interviewer; the other travels from the office to the interviewer and back to the office for control checking each time a segment is used for interviewing.

The two phase procedure has several advantages over a direct selection of scattered segments. Because the interviewer returns frequently to the same localities, it reduces chances for coverage errors. The sampling costs are spread over several surveys because virtually every segment in a chunk is selected for a study within the span of a year of two. It reduces variation in sample size, improves coverage (with the use of checks) and reduces the interviewers' problems at the time of interviewing.

ACCOUNTING AND CONTROL OF SAMPLE SIZE

The confusing discrepancies between sample addresses, dwelling units and respondents can be handled well only with thorough procedures. We developed a set of standard codes and a procedure for standard tabulations which are prepared after the completion of the field work for each study. From these we derive many useful estimates, the most important being:

$$\text{a) Coverage rate} = \frac{\text{occupied dwellings found}}{\text{occupied dwellings expected}}$$

$$\text{b) Response rate} = \frac{\text{interviews}}{\text{eligible respondents}} .$$

This information also allows us to plan the size of future studies more accurately. Furthermore, since information is available for individual places, the performance of one or two interviewers in a place can be followed and controlled better. During the interviewing period a continuous flow of completed questionnaires and control forms come into the central office where they are compared with a master control list for the study. Any questionable deviations from the master control are discussed with the interviewers, preferably while there is time to correct the field work.

EFFECTS OF NONCOVERAGE

It is useful to present a rough idea of the likely effect that the noncoverage has on some types of survey statistics. (The model is also relevant for nonresponses.) Denote by $\overline{Y}=Y/N=W_1\overline{Y}_1+W_2\overline{Y}_2$ the mean of some characteristic y of the population, where N is the number of population elements; W_1 and W_2 are respectively the proportions of coverage and noncoverage in the sample and $W_1+W_2=1$; also $\overline{Y}_1$ and $\overline{Y}_2$ are respectively, the means of the character-

istics in the two segments. Let us look at the relative bias (R.B.) in three important types of survey estimates.

a) The R.B. in estimating Y, the population total, by simple expansion is:

$$N(\bar{Y} - W_1\bar{Y}_1)/Y = N(W_2\bar{Y}_2)/N\bar{Y} = W_2\bar{Y}_2/\bar{Y} = W_2Z_a$$

Often $\bar{Y}_2 \doteq \bar{Y}$ and the R.B. is about the size of the noncoverage W_2; however, in situations where $\bar{Y}_2 \doteq 0$ the R.B. becomes negligible.

b) The R.B. in estimating the population mean $\bar{Y}$ is:

$$(\bar{Y} - \bar{Y}_1)/\bar{Y} = W_2(\bar{Y}_2 - \bar{Y}_1)/\bar{Y} = W_2Z_b$$

If $\bar{Y}_2 \doteq \bar{Y}_1$ the R.B. is negligible. For the bias to be important there must exist a coincidence of large noncoverage and large differences between the means of the two segments.

c) The difference $(\bar{Y}_A - \bar{Y}_B)$ between the means of two subclasses (A and B) of the population is subject to the R.B.:

$$\{[W_2(\bar{Y}_2 - \bar{Y}_1)]_A - [W_2(\bar{Y}_2 - \bar{Y}_1)]_B\}/(\bar{Y}_A - \bar{Y}_B)$$
$$\doteq W_2'\{(\bar{Y}_2 - \bar{Y}_1)_A - (\bar{Y}_2 - \bar{Y}_1)_B\}/(\bar{Y}_A - \bar{Y}_B) = W_2'Z_c$$

For the bias to be important in the comparisons of subclasses (such as age or occupation classes) the effect of the bias on one subclass must be much different than on the other subclass. (This assumes that the noncoverage in the two subclasses is about the same: $W_2' \doteq W_{2A} \doteq W_{2B}$).

The importance of a bias should be judged in relation to the other errors of the survey. Since nonresponse in our surveys is in the neighborhood of 10 per cent, the reduction of noncoverage from 10 to 3 per cent was important, but further reductions, while not negligible, are less so. Leaving aside the errors of response and of nonresponse, let us look at the relationship of noncoverage to the sampling error. This problem is broad and we shall have to reduce it to fit it into the modest ambitions of this paper. A simplified model can still be useful.

The sampling error can often be expressed in terms of the coefficient of variation (the ratio of the standard error to the estimate), denoted by $V = C/\sqrt{n}$. Here "C" is the coefficient of variation per element and n is the number of elements in a simple random sample. (Complex designs can be expressed in terms of "equivalent simple random size.") The relative bias can be expressed as R.B. $= W_2Z$; here Z is a generalized bias term representing, as required, any of the three terms in a) or b) or c) above; and for convenience we shall take Z always in the positive sense.

The ratio R.B./$V = W_2Z/V$ is the "effect" of the relative bias as a ratio of $V = C/\sqrt{n}$. It is also equal to the ratio of the bias to the standard error of the estimate. If R.B./$V = 0.1$ the $P = .05$ level of confidence is distorted to .0511; we call this effect (or less) "negligible." If R.B./$V = 0.2$, the $P = .05$ level is distorted to .0546; this effect is "small." If R.B./$V = 0.6$, the $P = .05$ level is distorted to .0921; this effect (or more) is "large." If $0.2 <$ R.B./$V < 0.6$, we say

the effect is "neither small nor large."[18] Table 524 shows the sizes of the generalized bias term (Z) necessary to produce a "negligible," (0.1), or a "small," (0.2), or a "large," (0.6), "effect" respectively, for different proportions of noncoverage (W_2) and for different sample sizes (n). Specifically, columns 2, 3 and 4 of Table 524 show the evaluation of

$$Z = \frac{C \text{ (Bias effect)}}{W_2\sqrt{n}}$$

for $C=1$ and for alternate bias effects.

The coefficient of variation (C) per element was assumed to be 1.0. This is in the nature of a central value. Income and liquid assets are in that neighborhood; for proportions near .50 to ($C=\sqrt{q/p}=\sqrt{.5/.5}$. For values of p near .20, $C=\sqrt{.8/.2}=2$. For some estimates C will be less.[19] When C is known to have a different value, it can be used to multiply the entries in Table 524.

Now for some rough conclusions.

a) For estimating Y, the value of Z is usually near 1.0. In Table 524 we see that if $W_2=.10$, "large" effects are produced for even the smallest samples. Even for $W_2=.03$ "large" effects are produced for all samples for $n \geq 400$. Because of this and because of the large sampling errors, estimates of Y by means of "simple expansions" of our samples—and of most household surveys—are not attempted. Our expansions use averages ($\overline{Y}$) and are based on Current Population Survey estimates; these correct for noncoverage by tying in to the decennial Census with current vital statistics.

b) Estimating values of $\overline{Y}$ are among the important objectives of our surveys. The sizes of our surveys are mostly near the 1600 to 2500 range—in "simple random equivalent size." There is some evidence from the Census Post Enumeration Survey[20] that for income Z_b it is about 0.2. In general, we would guess that for many items Z_b would be between 0.05 and 0.25. On that basis, when our noncoverage was $W_2=0.1$ most of our items had "effects" greater than "negligible," and some of them "large." Contrariwise, with our present W_2 about .03, we are in the region where none of our results suffers from a "large effect," and most are probably at or near negligible. For larger surveys ($n=10{,}000$) even this small noncoverage ($W_2=.03$) may result in "effects" above the "small" type.

c) For comparisons of subgroups of our surveys we are in the size classes of $n=100$ and $n=400$. (A comparison of two subgroups, each with $n=800$, has a sampling error about as large as a sample of $n=400$.) For this category we suspect that even the former large noncoverage ($W_2=.10$) usually resulted in small or negligible "effects." On the other hand, in larger samples these comparisons may also suffer effects which are not "small."

---

[18] For effect on confidence levels of the ratio Bias/Standard Error, see Table 1.1 in William G. Cochran, *Sampling Techniques*, John Wiley and Sons, New York: (1953), p. 9.

[19] For a display of values of C for some characteristics see Hansen, Hurwitz and Madow, *op. cit.*, pp 141–146.

[20] Monroe G. Sirken, E. Scott Maynes, and John A. Frechtling, "Reasons for differences in survey estimates of income: survey of consumer finances and post-enumeration survey," in *Report of the 1956 Conference on Research in Income and Wealth*, National Bureau of Economic Research, in press.

TABLE 524

VALUES OF Z NECESSARY TO PRODUCE A "NEGLIGIBLE," "SMALL," OR "LARGE" "BIAS EFFECT"; FOR DIFFERENT NONCOVERAGE RATES AND FOR ALTERNATE SAMPLE SIZES*

Sample Size	Noncoverage rate					
	$W_2 = 0.10$			$W_2 = 0.03$		
	Bias effect			Bias effect		
	0.1	0.2	0.6	0.1	0.2	0.6
n (1)	Negligible (2)	Small (3)	Large (4)	Negligible (2)	Small (3)	Large (4)
100	.100	.200	.600	.333	.667	2.000
400	.050	.100	.300	.167	.333	1.000
1,600	.025	.050	.150	.083	.167	.500
2,500	.020	.040	.120	.067	.133	.400
10,000	.010	.020	.060	.033	.067	.200

* The "*Bias Effect*" is the ratio of the noncoverage bias to the standard error; values of 0.1, 0.2 and 0.6 are denoted respectively by "negligible," "small," and "large."

Z is a generalized bias term in $Z = \dfrac{\text{C (Bias effect)}}{W_2\sqrt{n}}$; where W_2 is the proportion of noncoverage; and the sample size, n, is in "simple random equivalent." Z may refer to Z_a, Z_b or Z_c as explained in the text. It is assumed that $C = 1$.

COMPLETE COUNTERBALANCING OF IMMEDIATE SEQUENTIAL EFFECTS IN A LATIN SQUARE DESIGN

JAMES V. BRADLEY
Aero Medical Laboratory, Wright Air Development Center

If there is an even number of experimental conditions (Latin letters), it is possible to construct a Latin Square in which each condition is preceded by a different condition in every row (and in every column, if desired). These designs are useful in counterbalancing immediate sequential, or other order, effects. A simple, and easily remembered, procedure by which to construct such squares is described and illustrated. A proof is offered which shows that the procedure is valid for any size square having an even number of cells on a side.

1. INTRODUCTION

Suppose that each of $2k$ individuals is to be exposed once to each of $2k$ experimental conditions. In some cases it may be suspected that reaction to any given experimental condition will be influenced (a) by the *number* of experimental conditions preceding it, i.e., by its ordinal position in the sequence of presentation, (b) by the *particular* experimental condition *immediately* preceding it. This situation is commonly encountered in psychological experiments when every subject is run under all experimental conditions. Such experimental designs reduce variance at the expense of introducing sequential effects such as learning, fatigue and interactions between reactions to various experimental conditions, particularly those adjacent in order of presentation.

Under the conditions stated it is possible to counterbalance (a) and (b) above, i.e., for each of the $2k$ individuals a given experimental condition is preceded by a different number of conditions and by a different immediately preceding condition. The method has been discussed by Bugelski [1] and illustrated in the case where $2k=6$. A design appropriate to the case where $2k=4$ can be found in Cochran and Cox [2]. The general use of Latin Squares to eliminate unwanted experimental effects has been discussed by Grant [3]. It is the purpose of the present note to supplement this earlier material by presenting a very simple and easily remembered method by which such counterbalanced designs may be obtained and to prove its validity for any integral value of k.

2. METHOD

(a) Construct a $2k \times 2k$ table in which rows represent individuals and columns, from left to right, represent ordinal positions, from first to last, in which experimental conditions may occur in the sequence of condition-presentation. (b) Number the $2k$ experimental conditions successively from 1 to n. (c) Assign successively the integers from 1 to n to the n cells in the first row by proceeding from left to right entering only cells in odd-numbered columns, then reversing direction and filling in cells in even-numbered columns. (d) In each column,

starting with the number already entered in the top cell, proceed downward entering in each cell the integer immediately following the one in the cell above it, except that the integer n is considered to be followed by the integer 1.

It will be found that any given integer appears once in each row, once in each column, and is preceded once (and is followed once) by each of the other integers and once by none of them. The following table illustrates the method for the case where $2k=8$.

It is to be noted that while immediate sequential effects are completely counterbalanced, remote effects are not. For example in the table shown, the integer 1 is thrice preceded by an 8 and thrice by a 2 in the *second* cell preceding the cell which it occupies.

TABLE 526

ILLUSTRATION OF METHOD FOR CASE WHERE $2k=8$

Individual	Order of Presentation of Experimental Conditions (Cell Entries) to Individuals							
	Ist	IInd	IIIrd	IVth	Vth	VIth	VIIth	VIIIth
A	1	8	2	7	3	6	4	5
B	2	1	3	8	4	7	5	6
C	3	2	4	1	5	8	6	7
D	4	3	5	2	6	1	7	8
E	5	4	6	3	7	2	8	1
F	6	5	7	4	8	3	1	2
G	7	6	8	5	1	4	2	3
H	8	7	1	6	2	5	3	4

3. PROOF

The proof will consist merely of showing that the steps outlined under Procedure are sufficient to insure that each integer (a) appears exactly once in each column, (b) appears exactly once in each row, (c) is preceded exactly once by each of the other integers. Consider the integers from 1 to n arranged clockwise around a circle in increasing order of magnitude. Call the number of integers passed or reached in proceeding clockwise from one integer to another the "separation" between the two integers. By construction each cell entry in each column has a separation of 1 from the entry in the cell above it. Since there are n such cells in each column, each of the n integers appears exactly once in each column. And since, by construction, the top cell in each of the n columns is occupied by a different one of the n integers, each row must contain each integer exactly once. Phrased differently, each cell entry in a given row has the same separation from the corresponding cell entry in row 1; and since each of the n integers appears exactly once in row 1, they must do likewise in all other rows.

If there are an even number of columns, the cell entries in the first row will be

$$1, n, 2, n-1, 3, n-2, 4, \cdots, \frac{n}{2}-1, \frac{n}{2}+2, \frac{n}{2}, \frac{n}{2}+1,$$

in order from left to right. Proceeding in the same order, the separations between entries in pairs of adjacent cells are respectively $n-1$, 2, $n-3$, 4, $n-5$, 6, $\cdots$, 3, $n-2$, 1. Thus, the separations, in the order given, form a decreasing progression of odd numbers interlaced with an increasing progression of even numbers, and each of the $n-1$ separations appears exactly once.

By construction, each cell entry in a given row has a separation of exactly 1 from the corresponding cell entry in the row above it. Therefore every row will have the same separations between entries in adjacent cells as exist for the corresponding cells in row 1. That is to say the separation existing between a pair of adjacent cells in row 1 is an attribute of the corresponding pair of adjacent columns.

Now consider any given integer, i, and let p represent whatever integer immediately precedes i in the same row when i appears in the $2k \times 2k$ table. Since each of the $n-1$ pairs of adjacent columns is associated with a different one of the $n-1$ possible separations, and since i appears exactly once in each column, p must assume each of the $n-1$ integral values other than i exactly once. The same reasoning demonstrates that each of the other integers must follow i exactly once.

4. MODIFICATIONS

As the table stands, each condition is immediately preceded once by each of the $2k-1$ other conditions and once, when it appears in the Ist column, by no condition. Thus the set of immediately preceding conditions is not the same for every condition since a condition does not immediately precede itself. A remedy, suggested by the Referee, would be to add a zeroth column, identical to and just to the left of the Ist column. The data collected under the zeroth presentation would be discarded, and the sole function of the column would be to complete the balance by substituting "preceded by itself" for "preceded by nothing." Thus every condition would have exactly the same set of immediately preceding conditions. In many fields an experimenter might be reluctant to take data which subsequently must be discarded. In psychological experiments, however, since initial performance is frequently characterized by atypical scores and high variability, a "practice" period is often introduced previous to experimentation. The zeroth presentation therefore could serve the function of the practice period with no loss of research efficiency.

It has been pointed out by the Referee that the design property shown to exist within rows, i.e., that a given integer is immediately preceded (and is immediately followed) once by every other integer, can be made two-dimensional so that it is true of both rows and columns. The proof already given is valid for any permutation of rows. Therefore in order to counterbalance for adjacency in both dimensions, one need only construct the table already outlined, then permute rows so that reading downward in the first column one obtains the same sequence of cell entries as exist from left to right in the first row of the table already discussed. The table shown below represents this modification applied to the table presented earlier.

TABLE 528

MODIFICATION OF TABLE 1 COUNTERBALANCING ORDER EFFECTS IN COLUMNS AS WELL AS ROWS

	I	II	III	IV	V	VI	VII	VIII
A	1	8	2	7	3	6	4	5
H	8	7	1	6	2	5	3	4
B	2	1	3	8	4	7	5	6
G	7	6	8	5	1	4	2	3
C	3	2	4	1	5	8	6	7
F	6	5	7	4	8	3	1	2
D	4	3	5	2	6	1	7	8
E	5	4	6	3	7	2	8	1

It should be noted that while rowwise and columnwise adjacency is completely counterbalanced, adjacency in the diagonal dimension is not.

REFERENCES

[1] Bugelski, B. R., "A note on Grant's discussion of the Latin Square principle in the design of experiments," *Psychological Bulletin*, 46 (1949), 49–50.

[2] Cochran, W. G., and Cox, Gertrude, *Experimental Designs*, New York: Wiley, 1950, 103–5.

[3] Grant, D. A., "The Latin Square principle in the design and analysis of psychological experiments," *Psychological Bulletin*, 45 (1948), 427–42.

FACTORIAL EXPERIMENTATION IN SCHEFFÉ'S ANALYSIS OF VARIANCE FOR PAIRED COMPARISONS

OTTO DYKSTRA, JR.
General Foods Research Center

In applying paired comparisons methods we sometimes find that the samples to be compared represent factorial combinations. It is found that Scheffé's method of paired comparisons may be applied for an underlying factorial design. The larger designs tend to become prohibitive, since to double the number of test samples requires about four times the number of judgments. A blocking procedure is described where only twice the number of judgments are required to double the number of samples, taking in account an underlying 2-series factorial. The larger designs provide many degrees of freedom for order effects but may not provide any for a true error estimate. We obtain degrees of freedom for error by running some pairs twice in only one order. We again may use the underlying factorial to decide which pairs to run once in each order and which pairs to run twice in only one order.

1. INTRODUCTION

PAIRED comparisons techniques are finding wide application in industry, due in part to the recent literature on the subject. Although much of this literature deals with rank analysis, we shall concern ourselves with the analysis of variance of scored preferences as devised by Scheffé [4].

The purposes of this paper are four-fold:

1. To review the experimental setup and the mathematics of the Scheffé procedure (Section 2).
2. To discuss the application to factorials, with emphasis on the 2^{p-q} fractional factorials (Section 3).
3. To describe a blocking procedure which simultaneously fractionates the required numbers of observations (Section 4).
4. To describe a method for obtaining an error term in large experiments where no error term could otherwise be obtained (Section 5).

The procedures are illustrated by a worked example (Section 6).

The reader is referred to Daniel [1] and Davies [2] for a discussion of fractional factorials. The paired comparisons methods were recently discussed by Jackson and Fleckenstein [3].

2. SUMMARY OF SCHEFFE'S PROCEDURE

There are m treatments to be compared in all $M = m(m-1)/2$ possible pairs. We obtain $2r$ observations per pair, r in each order of presentation. The observations are scored statements of preference between pair members and range on a 7-point scale between $+3$ and -3. The preference scores are positive if the

first member of a pair is preferred, negative if the second member is preferred, and 0 if there is no preference. The points of the scale are described by Scheffé as:

Score	Statement on Pair (i, j)
3	I prefer i to j strongly
2	I prefer i to j moderately
1	I prefer i to j slightly
0	No preference
−1	I prefer j to i slightly
−2	I prefer j to i moderately
−3	I prefer j to i strongly

Considerable difficulty with approximate normality of observed scores was encountered when Scheffé's method was first applied, but by experimenting with our interviewing technique and the wording of the preference questions, we were able to find a technique which yielded nicely distributed results. It should be noted that 5-point scales may also be used. Again, it would be necessary to establish a proper testing technique before using these scales.

We define x_{ijk} as the kth observation on the ordered pair (i, j) and assume all x_{ijk} are independent random variables with mean μ_{ij} and variance σ^2.

The mean preference for treatment i to treatment j is μ_{ij} when presented in the order (i, j) and $-\mu_{ji}$ in the order (j, i). The average preference π_{ij} and the average difference due to order of presentation δ_{ij} are given by:

$$\begin{aligned} \pi_{ij} &= \tfrac{1}{2}(\mu_{ij} - \mu_{ji}), \qquad \pi_{ji} = -\pi_{ij}, \\ \delta_{ij} &= \tfrac{1}{2}(\mu_{ij} + \mu_{ji}), \qquad \delta_{ji} = \delta_{ij}. \end{aligned} \tag{1}$$

The average order effect δ is given by

$$\begin{aligned} \delta &= \sum_{i \neq j} \mu_{ij}/(2M) \\ &= \sum_{i<j} \delta_{ij}/M \end{aligned} \tag{2}$$

where M equals the number of pairs, $M = m(m-1)/2$. The first summation in (2), and in similar subsequent summations, is over all values of i and j with $i, j = 1, \cdots, m$ and $i \neq j$.

Scheffé, in his "hypothesis of subtractivity," postulates that there exist parameters $\alpha_1, \alpha_2, \cdots, \alpha_m$ characterizing the m treatments, such that the average preference π_{ij} for i to j is the difference between the corresponding parameters

$$\pi_{ij} = \alpha_i - \alpha_j, \qquad \pi_{ii} = 0, \tag{3}$$

so the "deviations from subractivity" γ_{ij} are given by

$$\gamma_{ij} = \pi_{ij} - (\alpha_i - \alpha_j). \tag{4}$$

Since $\sum_{j=1}^{m} \pi_{ij} = m\alpha_i - \sum_{j=1}^{m} \alpha_j$, Scheffé adds the assumption $\sum_{i=1}^{m} \alpha_i = 0$ to permit estimation of the α_i. Accordingly, $\sum_{j=1}^{m} \pi_{ij} = m\alpha_i$. The α_i may be regarded as main effects.

Since the x_{ijk} are random variables, e_{ijk} is used as the "error" in x_{ijk}, so that $e_{ijk}=x_{ijk}-\mu_{ij}$. The components which have been discussed are summarized below, together with the appropriate sums of squares resulting from partitioning the total sum of squares, using the analysis of variance table suggested by Scheffé. The degrees of freedom and sum of squares for order effects may be partitioned into 1 and $M-1$ degrees of freedom, as shown. As a result, we do not include in the totals the parenthesized entries M and S_δ. This table also shows the identities between the sums of squares. To complete the analysis of variance we compute the mean squares and the F-ratios, comparing each mean square to the error mean square. The last column has been included to show what the F-ratios are testing.

ANALYSIS OF VARIANCE

Source	Degrees of Freedom	Sum of Squares	H_0:
Main effects $(\alpha_i-\alpha_j)$	$m-1$	S_α	$\alpha_i=0$
Deviations from subtractivity (γ_{ij})	$M-(m-1)$	S_γ	$\gamma_{ij}=0$
Average preferences (π_{ij})	M	S_π	
Order effects (δ_{ij})	(M)	(S_δ)	
Average order effect (δ)	1	$2rM\delta^2$	$\delta=0$
Diff. among order effects $(\delta_{ij}-\delta)$	$M-1$	$S_\delta-2rM\delta^2$	$\delta_{ij}-\delta=0$
Means (μ_{ij})	$2M$	S_μ	
Error (e_{ijk})	$2M(r-1)$	S_e	
Total (x_{ijk})	$2rM$	S_t	

The following formulas will aid in estimating the various effects and computing the sums of squares:

Effect	*Estimation*	*Sums of Squares*
x_{ijk}		$S_t=\sum_{i\neq j}\sum_{k}x_{ijk}^2$
μ_{ij}	$\widehat{r\mu}_{ij}=\sum_{k=1}^{r}x_{ijk}$	$S_\mu=\frac{1}{r}\sum_{i\neq j}(r\mu_{ij})^2$
π_{ij}	$2\widehat{r\pi}_{ij}=\widehat{r\mu}_{ij}-\widehat{r\mu}_{ji}$	$S_\pi=\frac{1}{2r}\sum_{i<j}(2r\widehat{\pi}_{ij})^2$
δ_{ij}	$2\widehat{r\delta}_{ij}=\widehat{r\mu}_{ij}+\widehat{r\mu}_{ji}$	$S_\delta=\frac{1}{2r}\sum_{i<j}(2r\widehat{\delta}_{ij})^2$
δ	$2rM\widehat{\delta}=\sum_{i\neq j}\sum_{k}x_{ijk}$	$2rM\widehat{\delta}^2=\frac{1}{2rM}\left(\sum_{i\neq j}\sum_{k}x_{ijk}\right)^2$
α_i	$2rm\widehat{\alpha}_i=\sum_{j=i}^{m}2\widehat{r\pi}_{ij}$	$S_\alpha=\frac{1}{2rm}\sum_{i}(2rm\widehat{\alpha}_i)^2$

The advantage in using these formulas rather than those given by Scheffé arises from using integers throughout all the computations. The remainder of the sums of squares needed for the analysis of variance table are obtained by subtraction. Thus, $S_e = S_t - S_\mu$ and $S_\gamma = S_\tau - S_\alpha$.

3. APPLICATION TO FACTORIALS

Up to this point we have not discussed the m treatments, which may have been m brands of a product or m factorial combinations, in which case we may consider the $\hat{\alpha}_i$ as responses. Let us consider p factors, $A, B, \cdots, P$. The number of levels will be $a, b, \cdots, p$, respectively, so that there are $m = ab \cdots, p$ total combinations. Thus we may find $m-1$ orthogonal contrasts $\hat{\theta}_j$ such that

$$\hat{\theta}_j = \sum_{i=1}^{m} c_{ji}\hat{\alpha}_i, \qquad \sum_{i=1}^{m} c_{ji} = 0; \qquad j = 1, \cdots, m-1; \text{ and}$$
$$\sum_{i=1}^{m} c_{ji}c_{ki} = 0; \qquad j, k = 1, \cdots, m-1; j \neq k. \tag{5}$$

The sum of squares for main effects S_α is thus partitioned into $m-1$ components, each with 1 degree of freedom, with each contrast $\hat{\theta}_j$ contributing

$$2rm\hat{\theta}_j^2 \Big/ \sum_i c_{ji}^2 = \Big[\sum_i c_{ji}(2rm\hat{\alpha}_i)\Big]^2 \Big/ (2rm) \sum_i c_{ji}^2 \tag{6}$$

to S_α. In private correspondence Scheffé has shown

$$\text{Var}\,(\hat{\theta}_j) = \frac{\sigma^2}{2rm} \sum_i c_{ji}^2. \tag{7}$$

In the case of 2^p factorials, the above method is changed slightly if fractionation is necessary (we would usually run all 2^p combinations only if p is small, i.e. not larger than 4). For fractional factorials the number of combinations is $m = 2^{p-q}$, where 2^{-q} is the degree of fractionation. In determining the contrasts, $c_{ji} = \pm 1$, so that the sum of squares attributed to each contrast is

$$2r\hat{\theta}_j^2 = \Big[\sum_i c_{ji}(2rm\hat{\alpha}_i)\Big]^2 \Big/ m(2rm). \tag{8}$$

In practice, we would use the $2rm\hat{\alpha}_i$ as responses and use Yates' procedure to obtain the $m-1$ single degree of freedom contrasts. The variance of a factor effect $\hat{\theta}_j/m$ is given by

$$\text{Var}\,(\hat{\theta}_j/m) = \sigma^2/2rm^2. \tag{9}$$

This demonstrates the advantage in using a factorial design, since

$$\text{Var}\,(\hat{\alpha}_i) = (m-1)\sigma^2/2rm = (m-1)\,\text{Var}\,(\hat{\theta}_j/m). \tag{10}$$

The need for modification of the Scheffé procedure is demonstrated in Table 1OD. As m increases we need about four times as many observations to double the degrees of freedom for main effects. The blocking procedure described

below allows us to double the degrees of freedom for main effects by doubling the number of observations. Then the following section discusses a better allocation of the degrees of freedom for order effects and for error.

TABLE 533

ALLOCATION OF DEGREES OF FREEDOM ($r=1$)

Source	Degrees of Freedom			
	$m=16$	$m=32$	$m=64$	m
Main effects	15	31	63	$m-1$
Dev. subtractivity	105	465	1953	$(m-1)(m-2)/2$
Avg. preference	120	496	2016	$m(m-1)/2$
Order effects	120	496	2016	$m(m-1)/2$
Means	240	992	4032	$m(m-1)$
Error	0	0	0	0
Total	240	992	4032	$m(m-1)$

4. BLOCKING

Attention is now restricted to 2^{p-q} fractional factorials. We have 2^p possible factorial combinations. Assuming that 2^p is too large to run, we take a 2^{-q} fraction of the possible combinations, so that our design consists of 2^{p-q} factorial combinations. We shall run these combinations in two blocks with $m=2^{p-q-1}$ treatments per block. Each block is run as a separate design.[1] Thus all possible pairs of treatment combinations in each block are run, but the treatment combinations in one block are not compared with any of the treatment combinations in the other block. The following schematic diagram represents what is happening where $p-q=4$. The appropriate analysis procedure is to analyze each block separately and then combine the analysis of variance tables, adding the respective sums of squares and doubling the degrees of freedom.

	$1 \cdots 8$	$9 \cdots 16$
1	Block 1, $m=8$	Pairs not run
2 ⋮ 8		
9	Pairs not run	Block 2, $m=8$
⋮ 16		

[1] Using 2 blocks is similar to taking a 2^{-q-1} fraction in each block.

Since $r=1$ (we would prefer to run a bigger design rather than increase r), the variance of a factor averaged over both blocks is

$$V[(\hat{\theta}_j' + \hat{\theta}_j'')/2m] = \sigma^2/4m^2, \tag{11}$$

where $\hat{\theta}'_j$ and $\hat{\theta}''_j$ are the jth contrasts in the first and second blocks, respectively. As expected, this is twice what the variance would have been if the design had had $2m$ treatments, even though we have only one-half the number of judgments required for the usual $2m$ design.

5. ALLOCATION OF DEGREES OF FREEDOM FOR ORDER EFFECTS AND ERROR

When $r=1$ we have no degrees of freedom for error estimation. Scheffé considers this situation and suggests using the order effects mean square for the error term. An alternative suggested here is to run some pairs as indicated, once in each order, and to run some pairs twice but in only one order. We allow degrees of freedom for order effects in the former case and for error in the latter.

The procedure to be described is applied to the 2^p full factorial, to facilitate the description (there is no difficulty in applying the procedure to the 2^{p-q} fractional factorials, either unblocked or blocked as in Section 4). The m treatment combinations are divided into two groups, as in the blocking. For example, for five factors we may use ABCDE to group the treatment combinations. The two groups are shown in Table 534. Each treatment combination will be compared once in each order with each of the other treatment combinations within the same group. The treatment combinations in one group will be compared twice in only one order with each of the treatment combinations in the other group. Thus $m(m-2)/4$ pairs are run in each order and $m^2/4$ pairs in only one order.

TABLE 534

GROUPING FOR FIVE FACTORS, USING ABCDE

Group 1 Treatment Combinations				Group 2 Treatment Combinations			
(1)	ad	ae	de	a	d	e	ade
ab	bd	be	abde	b	abd	abe	bde
ac	cd	ce	acde	c	acd	ace	cde
bc	abcd	abce	bcde	abc	bcd	bce	abcde

In the event there is a bias for either the first or second sample presented, we must balance the direction of running the pairs for each treatment combination. An example of a balanced design for three factors is given in Table 535. The treatment combinations were grouped using ABC (it will generally be easiest to use the highest letter combination for grouping). Thus, those combinations which have 0 or 2 of the letters a, b, and c fall into the first group, while those which have 1 or 3 of the letters fall into the second group. The decision whether to run 0 or 2 judgments in the intergroup comparisons was based

TABLE 535

ALLOCATION OF JUDGMENTS (THE NUMBERS IN THE CELLS INDICATE THE NUMBERS OF JUDGMENTS TO BE OBTAINED IN THE INDICATED DIRECTIONS)

Treatment Combination Presented First		Treatment Combination Presented Second							
		Group 1				Group 2			
		(1)	ab	ac	bc	a	b	c	abc
Group 1	(1)	X	1	1	1	0	0	2	2
	ab	1	X	1	1	0	0	2	2
	ac	1	1	X	1	2	2	0	0
	bc	1	1	1	X	2	2	0	0
Group 2	a	2	2	0	0	X	1	1	1
	b	2	2	0	0	1	X	1	1
	c	0	0	2	2	1	1	X	1
	abc	0	0	2	2	1	1	1	X

on AB, which may be used regardless of the number of factors.[2] Where the treatment combinations in the first group are presented first in comparison with the treatment combinations in the second group, the comparisons of the combinations agreeing with respect to AB are run twice, while those differing are not run. (By "agreeing" we mean that both combinations have either 0 or 2 of the letters a and b, or that both have only one of the letters.) The situation is reversed when those in the second group are presented first against those in the first group, since the comparisons of the combinations differing with respect to AB are run twice. Thus, apart from the order of presentation, two judgments are obtained on each pairing. The degrees of freedom for experiments of various sizes, run in two blocks with the judgments arranged in each block as described above, are given in Table 536.

6. A WORKED EXAMPLE

The example chosen to illustrate the factorial, blocking, and balancing procedures was part of a series of experiments designed to improve an existing product. There were four factors in the experiment as described below:

Factor	Levels	Letters
Type of Raw Material	8	A, B, C
Acid Level	4	D, E
Acid Type	2	F
Additive Level	2	G

[2] In general, we should like to be clear of main effects in making this decision. Thus, we may use a letter combination with more letters, if the combination is not confounded with a main effect in a fractional design. It does seem, however, that the two letters are far enough removed from main effects for this purpose.

TABLE 536

ALLOCATION OF DEGREES OF FREEDOM, USING TWO BLOCKS OF m EACH AND $r=1$

Source	Degrees of Freedom			
	$2m=16$	$2m=32$	$2m=64$	$2m$
Main Effects	14	30	62	$2(m-1)$
Dev. Subtr.	42	210	930	$(m-1)(m-2)$
Avg. Preferences	56	240	992	$m(m-1)$
Order Effects	24	112	480	$m(m-2)/2$
Error	32	128	512	$m^2/2$
Total	112	480	1984	$2m(m-1)$

Admittedly any fraction of the 128 possible combinations would not allow the clear estimation of two-factor interactions. Nevertheless, only one-half of these combinations were run, with $I=-\overline{ABCDEFG}$ as the defining contrast. The blocking generator was $ABDF-\overline{CEG}$, giving 32 combinations in a block. ABCDE, representing the block live letters, was used to group the pairs, deciding which pairs to run once in each order and which to run twice in only one order. AB was used to decide in which of the two orders to run those pairs to be run in only one order.

TABLE 536A

EXPERIMENTAL TREATMENTS COMBINATIONS

Block 1				Block 2			
Group 1		Group 2		Group 1		Group 2	
1	(1)	17	af	1	fg	17	ag
2	ab	18	bf	2	abfg	18	bg
3	acfg	19	cg	3	ac	19	cf
4	bcfg	20	abcg	4	bc	20	abcf
5	ad	21	df	5	adfg	21	dg
6	bd	22	abdf	6	bdfg	22	abdg
7	cdfg	23	acdg	7	cd	23	acdf
8	abcdfg	24	bcdg	8	abcd	24	bcdf
9	aefg	25	eg	9	ae	25	ef
10	befg	26	abeg	10	be	26	abef
11	ce	27	acef	11	cefg	27	aceg
12	abce	28	bcef	12	abcefg	28	bceg
13	defg	29	adeg	13	de	29	adef
14	abdefg	30	bdeg	14	abde	30	bdef
15	acde	31	cdef	15	acdefg	31	cdeg
16	bcde	32	abcdef	16	bcdefg	32	abcdeg

TABLE 537

EXPERIMENTAL RESULTS FOR BLOCK 1

PART 1

Sample Tasted First (i)	Sample Tasted Second (j)															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		−2	−2	1	−3	2	2	2	2	2	2	−3	−3	2	3	−3
2	−3		1	3	3	0	2	0	2	3	−3	−2	2	2	3	2
3	−2	−1		1	−2	2	−1	0	2	0	−2	−2	−2	−2	2	2
4	−1	−2	−3		2	0	0	−3	3	3	1	−2	−2	0	−3	−2
5	0	−2	2	0		3	0	−3	−2	−3	−3	2	−2	−3	2	−3
6	−3	−3	−2	−1	3		2	−1	−3	1	−2	1	3	2	0	1
7	0	0	0	1	2	−2		−2	1	1	1	−2	2	−1	2	−2
8	0	1	−3	−3	2	−1	−3		0	3	2	1	3	−2	3	−2
9	−3	2	1	3	3	−3	2	3		3	0	2	−1	0	1	1
10	2	−1	−3	1	1	3	−3	1	−1		2	3	2	2	2	0
11	2	−2	3	2	2	−2	−3	2	0	−2		−3	−2	1	2	−3
12	−1	−3	−3	−2	1	2	1	2	1	0	−3		−3	3	−2	2
13	0	1	2	−1	1	3	0	0	2	−2	−2	3		−1	2	−2
14	0	−3	3	−3	−1	0	−2	1	0	2	2	3	−3		3	2
15	0	2	−3	1	−2	1	−1	−3	3	−3	−1	2	3	3		0
16	−3	2	2	−2	0	−1	0	−3	3	−3	−1	3	1	0	−2	

TABLE 537

EXPERIMENTAL RESULTS FOR BLOCK 1

PART 2

Sample Tasted First (i)	Sample Tasted Second (j)															
	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	—	—	2, 0	−3, −2	1, 2	0, 2	—	—	0, 1	1, 1	—	—	—	—	0, 1	2, −2
2	—	—	2, 2	−3, 1	2, 0	−1, 1	—	—	3, −2	0, 1	—	—	—	—	1, −3	−1, 2
3	2, 2	2, −3	—	—	—	—	−2, 2	−2, 2	—	—	0, −3	3, −2	3, −2	1, 3	—	—
4	−3, 1	0, 3	—	—	—	—	−2, −3	−3, 3	—	—	1, 2	2, −2	−2, 2	−2, 1	—	—
5	−2, 1	−3, −3	—	—	—	—	−3, −1	0, −2	—	—	2, 3	2, 2	−2, 0	−3, 2	—	—
6	2, 1	1, 0	—	—	—	—	0, −2	−2, −3	—	—	2, −3	−2, 1	1, −1	1, −3	—	—
7	—	—	0, −1	0, 2	3, −1	2, 3	—	—	2, 0	1, 1	—	—	—	—	0, −2	3, −2
8	—	—	0, −1	0, 0	0, 3	1, −3	—	—	−2, 1	−3, −2	—	—	—	—	−1, −1	2, 3
9	0, −3	−3, 3	—	—	—	—	−2, −2	3, 3	—	—	2, −3	−1, −2	3, 2	3, −2	—	—
10	3, −2	0, −2	—	—	—	—	−3, −3	−2, −2	—	—	0, 0	2, 2	−2, 1	−1, −2	—	—
11	—	—	−1, −2	0, −2	3, −1	2, 0	—	—	−1, −1	0, 0	—	—	—	—	3, −2	2, −2,
12	—	—	1, 3	0, 1	−3, −2	−2, 2	—	—	2, −1	2, 2	—	—	—	—	−3, −2	−2, 0
13	—	—	0, 2	−1, 2	0, 2	0, 0	—	—	2, 0	0, 1	—	—	—	—	0, 1	2, 1
14	—	—	0, 1	1, −2	2, −3	3, −3	—	—	−2, 1	−2, 3	—	—	—	—	3, −2	0, 0
15	3, −2	−2, 1	—	—	—	—	−3, −2	0, −1	—	—	0, −3	2, 2	0, −1	1, 3	—	—
16	−2, −3	−1, −2	—	—	—	—	−2, 2	−2, −1	—	—	2, 0	2, −3	−2, 3	−2, −2	—	—

TABLE 538

EXPERIMENTAL RESULTS FOR BLOCK 1

PART 1

Sample Tasted First (i)	Sample Tasted Second (j)															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
17	−3, −2	2, 0	—	—	—	—	1, −1	3, −2	—	—	−1, 3	2, 0	0, −3	2, 2	—	—
18	1, 3	2, −3	—	—	—	—	0, 3	−3, −2	—	—	−2, 1	3, 3	1, −3	−3, 3	—	—
19	—	—	−1, 3	2, −3	−2, 1	0, −3	—	—	3, 3	−3, 1	—	—	—	—	2, −3	1, −2
20	—	—	0, 3	2, −3	0, 3	3, 2	—	—	−2, −1	2, 2	—	—	—	—	3, 2	0, −2
21	—	—	−2, 1	−2, 1	−3, −1	−2, 0	—	—	2, −2	2, 3	—	—	—	—	3, 3	−3, 2
22	—	—	1, −2	2, −1	0, 3	0, 1	—	—	3, 2	3, 2	—	—	—	—	1, −3	2, 0
23	2, −2	1, −1	—	—	—	—	−1, 3	3, −1	—	—	−1, 1	2, −2	2, −2	2, −2	—	—
24	2, 2	−1, 2	—	—	—	—	−2, −3	−2, −1	—	—	−2, 3	1, 1	3, 2	−1, 3	—	—
25	—	—	3, −3	−1, −3	1, 3	−3, 2	—	—	−2, 0	−2, 0	—	—	—	—	−3, −3	2, 3
26	—	—	1, 1	1, 2	2, 3	−2, 0	—	—	−2, 2	−1, 0	—	—	—	—	2, 2	2, 2
27	0, 0	2, 3	—	—	—	—	3, −3	−3, 3	—	—	−1, 2	0, −2	0, 2	1, −3	—	—
28	3, −2	−2, −1	—	—	—	—	0, 1	2, 2	—	—	−1, −3	−3, 2	2, −3	2, 2	—	—
29	2, −2	3, 1	—	—	—	—	2, −1	−2, −2	—	—	1, −2	2, −3	2, 2	−2, 2	—	—
30	−2, −1	1, 0	—	—	—	—	3, −2	−1, −2	—	—	−1, 1	−3, −3	−2, 0	0, 1	—	—
31	—	—	3, −3	3, −1	2, 0	0, 0	—	—	2, 0	−2, −2	—	—	—	—	−1, 2	−1, −2
32	—	—	1, 0	−1, 2	2, −1	0, −3	—	—	−2, −3	3, −1	—	—	—	—	2, −2	3, 2

TABLE 538

EXPERIMENTAL RESULTS FOR BLOCK 1

PART 2

Sample Tasted First (i)	Sample Tasted Second (j)															
	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
17		1	3	−2	3	2	0	−3	−1	2	−1	3	−1	3	3	−2
18	0		−3	−3	−3	−3	−3	0	2	0	3	2	2	3	−2	1
19	2	0		3	−3	0	1	−1	2	−3	2	−2	3	−2	2	−1
20	1	2	0		−3	0	0	1	−2	3	−1	1	1	2	1	0
21	2	1	2	−2		−3	−3	2	3	−3	2	−3	3	−1	0	−2
22	−3	−1	−2	0	1		−3	−2	−3	3	3	2	−3	2	2	−1
23	1	−1	−3	−3	2	3		3	3	−1	−2	3	1	3	1	−1
24	−1	−2	1	2	−1	−3	−2		−3	2	−1	2	3	−1	−3	0
25	3	2	0	2	−3	−3	2	2		0	1	2	2	2	2	−2
26	3	2	−2	0	0	−2	−2	3	0		3	3	3	−3	1	2
27	2	−3	−3	−2	−3	−2	3	−2	3	−3		3	3	−1	0	3
28	−3	1	−1	3	−1	−3	2	−3	0	1	1		2	−2	−3	−2
29	2	−3	2	0	−2	−3	−3	−3	−1	−2	−1	2		−1	−2	−2
30	3	1	−3	2	−3	−2	1	1	−3	2	−3	1	1		2	−3
31	2	−3	2	0	−2	−2	−2	2	3	2	−2	2	2	1		3
32	−3	1	1	−3	2	0	3	0	3	0	1	2	0	−1	3	

TABLE 539

EXPERIMENTAL RESULTS FOR BLOCK 2

PART 1

Sample Tasted First (i)	Sample Tasted Second (j)															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		−3	2	−2	0	0	3	3	2	−2	3	3	1	1	2	3
2	−3		−2	3	−2	−2	−1	−2	3	−1	−3	2	2	−2	3	3
3	−3	2		0	−2	−1	3	2	−2	3	−3	−2	−3	−3	3	1
4	1	−3	3		−3	−1	0	−1	−2	−1	1	−1	2	−2	−3	2
5	−2	−1	3	−3		3	1	0	−2	3	−3	−2	2	3	2	3
6	3	3	0	0	3		−3	0	1	−2	2	2	3	1	2	1
7	1	2	−1	3	−2	−2		−3	2	−2	2	3	1	2	−1	−3
8	−3	0	−3	−3	−2	0	−2		−2	−1	1	−3	2	−3	−3	3
9	−2	1	3	−1	−3	−2	−3	2		0	−3	−3	1	0	−3	−3
10	−2	0	0	2	−3	2	−1	2	−1		−1	−2	2	2	2	1
11	3	3	−1	−1	2	−3	3	−2	0	−3		−3	2	3	−2	3
12	1	3	2	3	−2	−3	0	−2	−2	1	−3		0	−3	−1	2
13	−1	2	−2	−2	−2	0	0	−2	−2	2	0	2		−3	−2	0
14	−2	−2	−2	−2	−3	3	−2	−2	0	0	3	−2	0		−3	−2
15	−1	2	1	2	0	2	−2	−1	−2	2	3	2	3	3		2
16	3	−1	2	−2	−1	0	−3	−1	1	0	1	2	1	2	0	

TABLE 539

EXPERIMENTAL RESULTS FOR BLOCK 2

PART 2

Sample Tasted First (i)	Sample Tasted Second (j)															
	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
1	—	—	2, −3	2, 0	2, 1	−1, 2	—	—	−2, −3	2, 1	—	—	—	—	2, 2	2, −2
2	—	—	−3, 3	3, 3	3, 2	−3, −3	—	—	−2, 2	0, 0	—	—	—	—	3, 3	−2, −2
3	−2, −1	0, 0	—	—	—	—	0, −2	2, 0	—	—	−2, 2	−1, −3	2, 2	−2, −2	—	—
4	0, 3	0, 1	—	—	—	—	0, 3	0, −3	—	—	−3, −3	3, 3	2, 1	3, −1	—	—
5	3, 1	−3, 3	—	—	—	—	−3, −3	0, −2	—	—	3, 2	−2, 0	3, 1	3, −2	—	—
6	0, −2	−2, −3	—	—	—	—	2, −2	0, 0	—	—	3, 3	3, −1	2, −3	1, −2	—	—
7	—	—	3, 2	1, −2	−2, −1	0, 1	—	—	2, 3	−1, 3	—	—	—	—	−2, −2	2, 3
8	—	—	2, 2	−2, 2	−2, −2	2, 2	—	—	−1, 0	0, 3	—	—	—	—	1, 1	−3, −2
9	1, 3	2, −3	—	—	—	—	−2, −1	2, 2	—	—	0, 1	−2, 0	2, 1	1, 0	—	—
10	−1, −2	−2, −1	—	—	—	—	−1, 3	−2, −3	—	—	−2, 2	0, −1	−2, 2	2, −2	—	—
11	—	—	−2, 2	−1, −1	3, 2	1, 2	—	—	0, −2	2, 3	—	—	—	—	2, 2	2, −2
12	—	—	1, −2	−3, 3	1, 3	3, 3	—	—	−2, 3	1, 3	—	—	—	—	3, 0	2, 3
13	—	—	0, −1	2, −1	−1, −3	2, 0	—	—	1, −3	1, 3	—	—	—	—	2, 0	−3, −2
14	—	—	−3, −3	−3, −3	−1, −2	2, −2	—	—	3, 2	2, −3	—	—	—	—	−3, 0	1, 2
15	−1, 1	−1, 2	—	—	—	—	0, −3	0, −2	—	—	−2, 3	−2, −2	3, 0	2, −3	—	—
16	2, −1	2, −1	—	—	—	—	−3, −2	0, −1	—	—	1, 2	1, −1	0, −3	3, 2	—	—

TABLE 540
EXPERIMENTAL RESULTS FOR BLOCK 2
PART 1

Sample Tasted First (i)	Sample Tasted Second (j)															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
17	−2, −3	1, 3	—	—	—	—	0, 1	−1, −2	—	—	0, 2	−3, 3	3, −2	0, 1	—	—
18	3, −3	2, −1	—	—	—	—	0, 2	2, −2	—	—	2, −2	2, 3	3, 2	3, 3	—	—
19	—	—	0, −1	2, −2	1, −3	−3, 3	—	—	−2, 2	0, 0	—	—	—	—	2, −3	2, 3
20	—	—	3, 3	2, −2	−2, −1	2, −1	—	—	3, −1	0, 1	—	—	—	—	−1, −2	0, 2
21	—	—	−2, 2	−2, 0	−3, −2	1, 0	—	—	0, 2	3, 1	—	—	—	—	−2, 3	−3, 3
22	—	—	1, 2	−2, 3	−3, 2	−1, −3	—	—	−2, −3	2, −2	—	—	—	—	1, 2	1, −1
23	−3, 1	−1, 1	—	—	—	—	2, −2	−1, −2	—	—	2, 1	−3, 1	0, −2	2, −3	—	—
24	−1, 2	2, 0	—	—	—	—	2, 2	1, 3	—	—	0, 1	−2, 1	−2, 2	−2, 1	—	—
25	—	—	−1, −3	−2, −1	−1, 2	−3, −2	—	—	3, −3	2, −2	—	—	—	—	1, −3	−2, 1
26	—	—	−2, 3	2, −1	−1, −3	0, −2	—	—	0, 2	0, 3	—	—	—	—	1, −2	3, −3
27	3, 2	2, 0	—	—	—	—	1, −2	−2, 3	—	—	−2, 1	−2, 1	−1, −3	−3, 3	—	—
28	3, −2	0, −2	—	—	—	—	−3, 3	2, 2	—	—	0, 0	−2, 2	−2, 0	0, 0	—	—
29	2, −2	3, 3	—	—	—	—	−1, 3	0, −3	—	—	2, 2	−3, 2	−2, −3	3, 3	—	—
30	−1, −1	−1, 1	—	—	—	—	3, 0	−3, 2	—	—	2, 2	2, 2	1, 2	1, 2	—	—
31	—	—	−1, 1	3, −1	−1, −1	2, −3	—	—	2, 2	−3, −3	—	—	—	—	0, −2	0, 1
32	—	—	−3, 2	2, −2	−1, −1	−3, −2	—	—	−3, 3	−2, −3	—	—	—	—	0, −3	0, −2

TABLE 540
EXPERIMENTAL RESULTS FOR BLOCK 2
PART 2

Sample Tasted First (i)	Sample Tasted Second (j)															
	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
17		3	1	2	−2	2	−1	−1	2	1	1	−2	0	−1	−3	1
18	−2		−1	1	1	−3	−1	−2	3	−3	3	2	3	−1	−1	2
19	−1	−2		−2	−2	3	2	−2	2	0	−2	0	3	−3	2	2
20	−3	0	0		2	2	2	2	0	1	−2	2	−3	−3	2	−2
21	3	−1	−3	2		2	3	−2	1	1	3	2	−1	−1	−3	2
22	−2	0	3	3	0		−2	1	2	2	−3	−2	1	−3	2	−1
23	−1	−1	2	−3	0	−2		2	2	−3	1	1	3	1	2	−3
24	−1	−3	2	−2	−2	1	−1		0	2	3	1	−2	0	−3	−2
25	2	−2	1	2	1	2	2	2		0	1	−2	3	2	−1	3
26	−1	−3	0	−2	3	−3	−3	1	2		−1	0	1	1	2	2
27	0	−3	−2	−2	1	3	−1	−2	2	2		−2	−2	2	2	3
28	0	3	0	−3	0	−2	−1	−3	−2	−1	−3		0	2	−2	−1
29	−2	3	2	−1	1	2	−3	2	0	−2	3	−1		−1	−2	−1
30	−3	0	1	−3	−3	2	−2	−3	3	−2	2	2	1		2	−1
31	−2	0	−2	−2	2	−1	−2	−2	2	−3	−2	−2	−1	−2		1
32	−2	−1	−3	2	−2	−2	−1	−2	3	−1	3	0	2	−1	0	

TABLE 541

ANALYSIS OF VARIANCE ($2m=64$, $r=1$, 2 BLOCKS)

Source	d.f.	S.S.	M.S.	F
"Main Effects"	(62)	(427.47)		
ABC (Type of Raw Material)	7	22.48	3.21	.76
DE (Acid level)	3	93.80	31.27	7.43
F (Acid Type)	1	25.95	25.95	6.16
G (Additive Level)	1	14.30	14.30	3.40
DEF	3	81.04	27.01	6.42
Others	47	189.90	4.04	.96
Dev. Subtr.	930	3720.53	4.00	.95
Avg. Preferences	992	4148.00		
Order Effects	(480)	(2159.00)		
Avg. Order Effect	1	6.55	6.55	1.56
Resid. Order Effect	479	2152.45	4.49	1.07
Error	512	(2147.00)	4.21	
Total	1984	8464.00		

The treatment combinations are given in Table 536a, the experimental results in Tables 537–540, and the analysis of variance in Table 541. Significant effects were attributed to acid level, acid type, and the interaction of acid level and acid type. The averages for the eight combinations of acid type and level are given in Table 541A. It can be seen from the averages that preference reaches a maximum at the second level for Acid Type 1 and at the third level for Acid Type 2. With both types at their best levels, Acid Type 2 is more preferred than Acid Type 1.

TABLE 541A

AVERAGES FOR THE COMBINATIONS OF ACID TYPE AND LEVEL

Acid Type	Acid Level			
	1	2	3	4
1	−.053	.211	−.027	−.449
2	−.141	.066	.338	.055

The mean squares for the other factors are not large enough to be significant. The mean squares for the remainder of the interactions, for deviations from subtractivity, and for order effects agree well with the error.

7. REMAINING PROBLEMS

a. Blocking was done in two blocks. It may be necessary to carry out larger experiments in four blocks, which certainly would be advisable because of the waste of degrees of freedom on subtractivity, order effects, and error.

b. Balanced sub-sets. Since we have very high precision in the larger experiments, it may be desirable to omit some pairs in a block. The main consequence of this would be to reduce the degrees of freedom for subtractivity.

REFERENCES

[1] Daniel, C., "Fractional replication in industrial research." *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954.

[2] Davies, O. L., *The Design and Analysis of Industrial Experiments*. London: Oliver and Boyd, 1954.

[3] Jackson, J. E., and Fleckenstein, M., "An evaluation of some statistical techniques used in the analysis of paired comparison data," *Biometrics*, 13 (1957), 51–64.

[4] Scheffé, Henry, "An analysis of variance for paired comparisons," *Journal of the American Statistical Association*, 47 (1952), 381–400.

DESIGN AND OPERATION OF A DOUBLE-LIMIT VARIABLES SAMPLING PLAN

ACHESON J. DUNCAN
Johns Hopkins University

The paper presents and illustrates the use of three charts and a table for the design and operation of a double-limit variables sampling plan. Two of the charts are plots of the Lieberman-Resnikoff tables of the minimum variance unbiased estimate of the proportion of a normal universe exceeding specified limits given sample estimates of the mean and standard deviation or average range. The table is an extension of the Patnaik table of conversion factors for use of the average range in lieu of the standard deviation.

1. INTRODUCTION

LIEBERMAN and Resnikoff [2] have published a set of variables sampling plans that were based largely on Blackwell and Girshick's minimum-variance unbiased estimator of the fraction defective of a normal universe. The article contained a standard set of 16 plans and their operating characteristic curves, together with tables for use in operation of the plans. These plans are for the most part identical with those of the new variables standard of the Department of Defense [6], Mil. St. 414.

Whereas many will be willing to use one of the plans offered by Lieberman and Resnikoff [2], there may be others who would prefer to design a variables plan to meet their own special needs. For single limit plans, standard procedures have been known for some time for deriving a variables plan to pass through two points on an operating characteristic curve [7]. The special feature of interest in the work of Lieberman and Resnikoff is that they offer statistically efficient plans for situations in which there are double specification limits. It is the purpose of this article to present a set of charts and a table that will be helpful in deriving and operating a double-limit variables sampling plan of the Lieberman-Resnikoff type.

2. DESIGN AND OPERATION OF A DOUBLE-LIMIT SAMPLING PLAN BASED ON THE MEAN AND STANDARD DEVIATION OF THE SAMPLE

For the sake of brevity we shall call the Lieberman and Resnikoff type of plan a p^* plan. To design a p^* double-limit plan that will have a given p_1, p_2, α and β, where $1-\alpha$ is the probability of accepting lots of p_1 quality and β is the probability of accepting lots of p_2 quality, we take the following steps:

(a) We design a single-limit plan for these same p_1, p_2, α and β. For this the formulas offered by Wallis [7] may be used, viz.,[1]

[1] For a more precise method see [5].

$$k = \frac{K_\alpha K_2 + K_\beta K_1}{K_\alpha + K_\beta}$$
$$n = \frac{2(K_\alpha + K_\beta)^2 + (K_\alpha K_2 + K_\beta K_1)^2}{2(K_1 - K_2)^2} \tag{1}$$

where K_ϵ is defined by

$$\int_{K_\epsilon}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz = \epsilon \tag{2}$$

and $K_1 = K_{p_1}$ and $K_2 = K_{p_2}$.

(b) For the k and n so derived, we find β_{p^*} from the formula

$$\beta_{p^*} = \frac{1 - \frac{\sqrt{n}}{n-1} k}{2} \tag{3}$$

given by Lieberman and Resnikoff [2]. We then enter Fig. 545 with an abscissa value equal to β_{p^*} and for the computed value of n read or interpolate the value of p^*. Fig. 545 is merely a graph of the incomplete Beta function for equal parameters and was derived from Karl Pearson's *Tables of the Incomplete Beta Function*, taking $p^* = B(p, q)$, $\beta_{p^*} = x$ and $n = 2p + 2$ with $p = q$.

The two steps stated above will yield a double-limit p^* plan for specified p_1, p_2, α and β. This plan will be operated as follows. Take a random sample of size n and compute $\overline{X} = \sum X/n$ and

$$s = \sqrt{\frac{\sum (X - \overline{X})^2}{n - 1}}$$

Compute

$$C_L = \frac{\overline{X} - L}{s} \quad \text{and} \quad C_U = \frac{U - \overline{X}}{s},$$

where L and U are the lower and upper specification limits respectively, and from Fig. 546 determine $\hat{p}_L$ and $\hat{p}_U$. If $\hat{p}_L + \hat{p}_U \leq p^*$, accept the lot; otherwise reject it. Fig. 546 is merely a graphical representation of Lieberman and Resnikoff's Table IV. In most cases it can be read with sufficient accuracy for practical operation of a plan. If greater accuracy is desired, the table should be used.[2]

To illustrate the above procedure let us design a double-limit variables plan for which $p_1 = 0.014$ and $p_2 = 0.09$, $\alpha = 0.05$ and $\beta = 0.10$. We shall have from (1) and tables of the normal probability distribution

$$k = \frac{(1.6449)(1.3408) + (1.2816)(2.1973)}{(1.6449 + 1.2816)} = 1.716$$

[2] The graph has some theoretical interest, in that it affords a ready comparison between the minimum variance unbiased estimate of p and the estimate frequently made by treating $C_L = (\overline{X} - L)/s$, for example, as a normal deviate. It indicates that these estimates are approximately the same for $n > 100$, but differ markedly for small n. For low C_L, the minimum variance method yields a slightly higher estimate, but for larger C_L a much smaller estimate of p. The same statements, of course, pertain to C_U.

and

$$n = \frac{2(1.6449 + 1.2816)^2 + [(1.6449)(1.3408) + (1.2816)(2.1973)]^2}{2(2.1973 - 1.3408)^2}$$

$$= 28.86 \doteq 29.$$

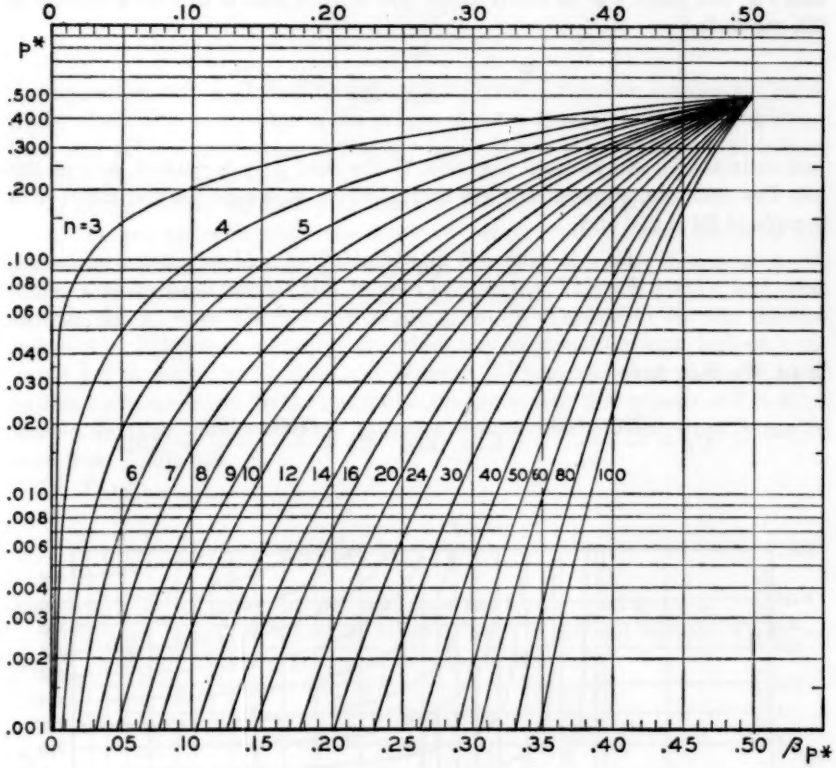


FIG. 545. Chart for determining p^* from β_{p^*}. (Graph of incomplete beta function, equal parameters.)

For average range plans:

$$\beta_{p^*} = \frac{1 - k/\sqrt{v}}{2}$$

For standard deviation plans:

$$\beta_{p^*} = \frac{1 - \dfrac{\sqrt{n}}{n-1}\, k}{2}$$

With $k=1.716$ and $n=29$ we have

$$\beta_{p^*} = \frac{1 - \frac{\sqrt{29}}{28}(1.716)}{2} = 0.335,$$

and Fig. 545 yields a p^* of about 0.040. The desired plan is to take a sample of 29, compute

$$C_L = \frac{\bar{X} - L}{s} \quad \text{and} \quad C_U = \frac{U - \bar{X}}{s},$$

and estimate $\hat{p}_L$ and $\hat{p}_U$ from Fig. 546. If the sum $\hat{p}_L + \hat{p}_U \leq 0.040$, accept the lot. For example, suppose $L=500$ and $U=700$, and suppose the mean of a sample of 29 is 570 and

$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{n-1}}$$

is 40. We then have

$$C_L = \frac{570 - 500}{40} = 1.75 \quad \text{and} \quad C_U = \frac{700 - 570}{40} = 3.25.$$

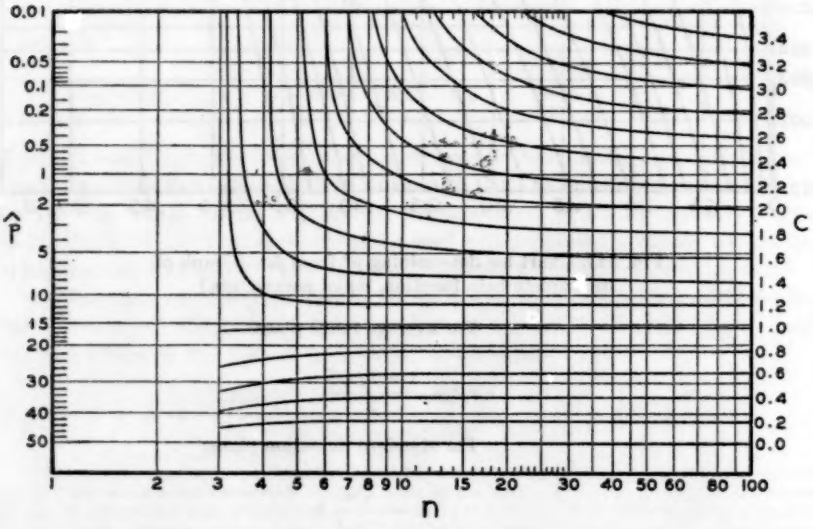


FIG. 546. Chart for estimating p from the quality index C and sample size n. Standard deviation plans. (Chart of Lieberman and Resnikoff [2], Table IV.)

Fig. 546 yields $\hat{p}_L \doteq 0.037$ and $\hat{p}_U \doteq 0.0001$. The sum $\hat{p}_L + \hat{P}_U = 0.0371$. Since this is less than 0.040, we accept the lot.

3. DESIGN AND OPERATION OF A DOUBLE-LIMIT SAMPLING PLAN BASED ON THE MEAN AND AVERAGE RANGE OF THE SAMPLE

For those who prefer to use the average range $\overline{R}$ instead of s, a similar procedure can be employed. Patnaik has shown [3] that

$$\frac{\nu(\overline{R}/d_2^*)^2}{\sigma^2}$$

has approximately a χ^2 distribution with degrees of freedom ν given by Table 1AD, where $\overline{R}$ is an average of the ranges of g sub-groups consisting of m items each and d_2^* is a constant given by the table. When $\overline{R}$ is to be used, therefore, we take the following steps:

(a) Same as Step (a) for a plan using s. (See Section 2 above.)

(b) To get approximately the same operating characteristic curve as obtained by using s, we look in Table 548 for the value of ν that is close to $n-1$, where n is the total sample size derived in Step (a). This will yield a new total sample size $n_R = gm$ where g and m are given by Table 548 for the derived value of ν. If there is a choice, we select the combination with highest ν for equal gm. A study by Grubbs and Weaver [1] suggests that when gm is divisible by 8, $m=8$ will be the optimum sub-group size. Sub-groups of 7 or 9 are almost as good, however. This optimum character of sub-group of 8 can be seen from Table 548.

(c) Take

$$\beta_{p^*} = \frac{1 - k/\sqrt{\nu}}{2}$$

and determine[3] p^* from Fig. 545 with n of the figure equal to $\nu+1$.

This range plan is operated as follows. Take a random sample of $gm = n_R$ items, subdivide it into g groups of m each. Compute

$$C_L = \frac{\overline{X} - L}{\overline{R}/d_2^*} \quad \text{and} \quad C_U = \frac{U - \overline{X}}{\overline{R}/d_2^*}$$

and from Fig. 549 find $\hat{p}_L$ and $\hat{p}_U$. Accept the lot if $\hat{p}_L + \hat{p}_U \leq p^*$; otherwise reject. Fig. 549 is a graphical representation of Lieberman and Resnikoff's Table VI.

To illustrate let us return to the example worked out in the previous section and derive an average range plan for the same values of p_1, p_2, α and β. These were $p_1 = 0.014$, $p_2 = 0.09$, $\alpha = 0.05$ and $\beta = 0.10$. In this case, $n-1 = 29-1 = 28$. In Table 548 the gm combination that yields a ν closest to 28 is $g=3$, $m=13$. The next step is to find p^*. We have

$$\beta_{p^*} = \frac{1 - 1.716/\sqrt{28}}{2} = 0.338.$$

[3] See [2]. This involves an additional approximation in that $\hat{p}\ (\overline{X}, \overline{R})$ is assumed to have approximately the same distribution as $\hat{p}\ (\overline{X}_{\nu+1}, s_\nu)$.

TABLE 548

VALUES ASSOCIATED WITH THE DISTRIBUTION OF THE AVERAGE RANGE[1]

$\nu(\overline{R}/d_2^*)^2/\sigma^2$ is distributed approximately as χ^2 with ν degrees of freedom; $\overline{R}$ is the average range of g subgroups, each of size (m)

Number of Samples (g)	Size of Samples (m)																							
	2		3		4		5		6		7		8		9		10		11		12		13		14		15	
	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*	ν	d_2^*
1	1.0	1.41	2.0	1.91	2.9	2.24	3.8	2.48	4.7	2.67	5.5	2.83	6.3	2.96	7.0	3.08	7.7	3.18	8.4	3.27	9.0	3.35	9.6	3.42	10.2	3.49	10.8	3.55
2	1.9	1.28	3.8	1.81	5.7	2.15	7.5	2.40	9.2	2.60	10.8	2.77	12.3	2.91	13.8	3.02	15.1	3.13	16.5	3.22	17.8	3.30	19.0	3.38	20.2	3.45	21.3	3.51
3	2.8	1.23	5.7	1.77	8.4	2.12	11.1	2.38	13.6	2.58	16.0	2.75	18.3	2.89	20.5	3.01	22.6	3.11	24.6	3.21	26.5	3.29	28.4	3.37	30.2	3.43	31.9	3.50
4	3.7	1.21	7.5	1.75	11.2	2.11	14.7	2.37	18.1	2.57	21.3	2.74	24.4	2.88	27.3	3.00	30.1	3.10	32.7	3.20	35.3	3.28	37.8	3.36	40.1	3.43	42.4	3.49
5	4.6	1.19	9.3	1.74	13.9	2.10	18.4	2.36	22.6	2.56	26.6	2.73	30.4	2.87	34.0	2.99	37.5	3.10	40.9	3.19	44.1	3.28	47.1	3.35	50.1	3.42	52.9	3.49
6	5.5	1.18	11.1	1.73	16.6	2.09	22.0	2.35	27.1	2.56	31.8	2.73	36.4	2.87	40.8	2.99	45.0	3.10	49.0	3.19	52.8	3.27	56.5	3.35	60.1	3.42	63.5	3.49
7	6.4	1.17	12.9	1.73	19.4	2.09	25.6	2.35	31.5	2.55	37.1	2.72	42.5	2.87	47.5	2.99	52.4	3.10	57.1	3.19	61.6	3.27	65.9	3.35	70.0	3.42	74.0	3.48
8	7.2	1.17	14.8	1.72	22.1	2.08	29.3	2.35	36.0	2.55	42.4	2.72	48.5	2.87	54.3	2.98	59.9	3.09	65.2	3.19	70.3	3.27	75.3	3.35	80.0	3.42	84.6	3.48
9	8.1	1.16	16.6	1.72	24.8	2.08	32.9	2.34	40.5	2.55	47.7	2.72	54.5	2.86	61.0	2.98	67.3	3.09	73.3	3.18	79.1	3.27	84.6	3.35	90.0	3.42	95.1	3.48
10	9.0	1.16	18.4	1.72	27.6	2.08	36.5	2.34	44.9	2.55	52.9	2.72	60.6	2.86	67.8	2.98	74.8	3.09	81.5	3.18	87.9	3.27	94.0	3.34	99.9	3.42	105.6	3.48
11	9.9	1.16	20.2	1.71	30.3	2.08	40.1	2.34	49.4	2.55	58.2	2.72	66.6	2.86	74.6	2.98	82.3	3.09	89.6	3.18	96.6	3.27	103.4	3.34	109.9	3.41	116.2	3.48
12	10.8	1.15	22.0	1.71	33.0	2.07	43.7	2.34	53.9	2.55	63.5	2.72	72.7	2.85	81.3	2.98	89.7	3.09	97.7	3.18	105.4	3.27	112.8	3.34	119.9	3.41	126.7	3.48
13	11.6	1.15	23.9	1.71	35.7	2.07	47.4	2.34	58.4	2.55	68.8	2.71	78.7	2.85	88.1	2.98	97.2	3.09	105.8	3.18	114.1	3.27	122.2	3.34	129.9	3.41	137.3	3.48
14	12.5	1.15	25.7	1.71	38.5	2.07	51.0	2.34	62.8	2.54	74.0	2.71	84.7	2.85	94.8	2.98	104.6	3.08	113.9	3.18	122.9	3.27	131.5	3.34	139.8	3.41	147.8	3.48
15	13.4	1.15	27.5	1.71	41.2	2.07	54.6	2.34	67.3	2.54	79.3	2.71	90.8	2.85	101.6	2.98	112.1	3.08	122.1	3.18	131.7	3.26	140.9	3.34	149.8	3.41	158.3	3.48
d_3		1.13		1.69		2.06		2.33		2.53		2.70		2.85		2.97		3.08		3.17		3.26		3.34		3.41		3.47
c.d.	0.88		1.82		2.74		3.62		4.47		5.27		6.03		6.76		7.45		8.12		8.76		9.38		9.97		10.54	

[1] See [3]. Also see *Biometrika Tables for Statisticians*, Vol. I, p. 178 and Acheson J. Duncan, "The use of ranges in comparing variabilities," *Industrial Quality Control*, Vol. XI, No. 5, February 1955. Table 1AD is an extension of these earlier tables. In general, the degrees of freedom will be given approximately by the reciprocal of $(-2+2\sqrt{1+2(\text{c.v.})^2/g})$ where c.v. is the coefficient of variation (d_3/d_2) of the range and g is the number of subgroups. Also d_2^* is given approximately by d_2 (i.e. the infinity value of d_2^*) times $(1+1/4\nu)$. Cf. [3]. Values of ν are also very readily built up from the constant differences. Table 1AD is a basic table that may be used whenever the average range is used in lieu of s.

Fig. 545 then yields for $\beta_{p^*}=0.338$ and $n=29$, a $p^* \doteq 0.042$.

Our plan will operate as follows. Take a random sample of 39 items and subdivide it into 3 sub-groups of 13. For each sub-group find the range R and compute $\overline{R}$. Compute

$$C_L = \frac{\overline{X} - L}{\overline{R}/d_2^*} = \frac{\overline{X} - 500}{\overline{R}/3.37} \quad \text{and} \quad C_U = \frac{U - \overline{X}}{\overline{R}/d_2^*} = \frac{700 - \overline{X}}{\overline{R}/3.37}.$$

From Fig. 549 estimate $\hat{p}_L$ and $\hat{p}_U$. If $\hat{p}_L + \hat{p}_U \leq 0.042$, accept the lot.

4. OPERATING CHARACTERISTIC CURVES FOR DOUBLE-LIMIT VARIABLES PLANS

The operating characteristic curve for a double-limit plan using s (or $\overline{R}$ adjusted for sample size) will be approximately that for the single-limit plan using s. (See [2].) Wallis's formulas for this [7] are:[4]

$$\text{Probability of Acceptance } \phi = \int_{K_\phi}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\, dz$$

where

$$K_\phi = \frac{k - K_p}{\sqrt{\dfrac{1}{n} + \dfrac{k^2}{2(n-1)}}}$$

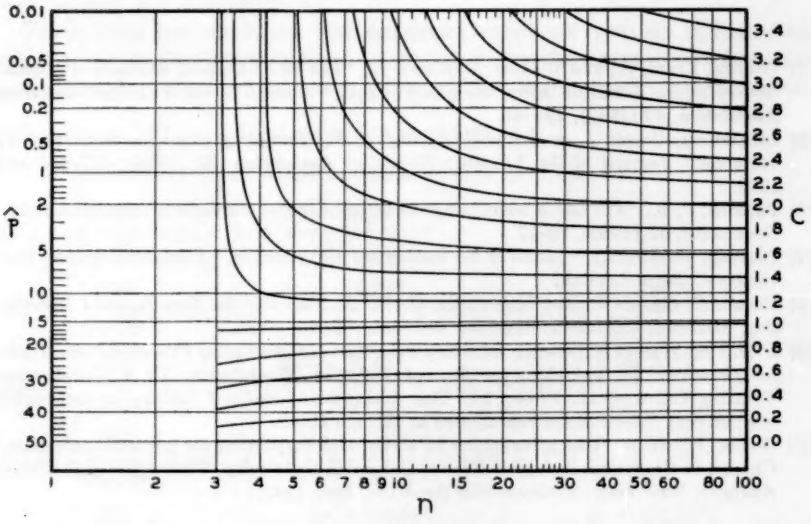


FIG. 549. Chart for estimating p from quality index C and sample size n. Average range plans. (Lieberman and Resnikoff [2], Table VI.)

[4] More precise results would be given by tables of the non-central t distribution. See [5].

and K_p is given by

$$\int_{K_p}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\, dz = p.$$

For example, if $p = 0.05$, $K_p = 1.645$ and under the variables plan derived in Section 2, the probability of accepting lots of 0.05 quality would be the probability that a standard normal deviate should exceed

$$\frac{k - K_p}{\sqrt{\frac{1}{n} + \frac{k^2}{2(n-1)}}} = \frac{1.716 - 1.645}{\sqrt{\frac{1}{29} + \frac{(1.716)^2}{2(28)}}} = \frac{0.071}{0.295} = 0.241$$

which is given by the normal tables as 0.40.

5. A CAUTION

The reader should be careful to note that these variables plans are all based on the assumption of normally distributed lots. We may state somewhat arbitrarily that as long as $p_1 \geq 0.01$, little error is likely to result if lots are moderately skewed or leptokurtic. For a p_1 much less than 0.01, however, serious doubts arise as to the validity of the operating characteristic curves derived on the assumption of normality. Care must also be exercised in rejecting lots under a variables plan that have been screened by the supplier and are thus truncated at one or both ends. Under such circumstances lots are usually submitted to a supplementary attributes plan before final rejection is allowed [6].

REFERENCES

[1] Grubbs, Frank E., and Weaver, Chalmers L., "The best unbiased estimate of population standard deviation based on group ranges," *Journal of the American Statistical Association*, 42 (1947), 224–41.

[2] Lieberman, Gerald J., and Resnikoff, George J., "Sampling plans for inspection by variables," *Journal of the American Statistical Association*, 50 (1955), 457–516 and 1333.

[3] Patnaik, P. B., "The use of mean range as an estimator of variance in statistical tests," *Biometrika*, 37 (1950), 78–87.

[4] Pearson, Karl (Ed.), "*Tables of the incomplete beta function*," *Biometrika* Office, University College, London.

[5] Resnikoff, George J., and Lieberman, Gerald J., *Tables of the Non-Central t Distribution*. Stanford, California: Stanford University Press, 1957.

[6] U. S. Department of Defense. Military Standard 414, *Sampling Procedures and Tables for Inspection by Variables for Percent Defective*. Washington: U. S. Government Printing Office, 1957. Price 65¢. The Lieberman-Resnikoff tables for estimating the per cent defective are reproduced in this standard.

[7] Wallis, W. Allen, "Use of variables in acceptance inspection for per cent defective," Chap. 1 in Statistical Research Group, Columbia University, *Techniques of Statistical Analysis*. New York: McGraw-Hill Book Co., Inc., 1947.

EMPIRIC INVESTIGATION OF A TEST OF HOMOGENEITY FOR POPULATIONS COMPOSED OF NORMAL DISTRIBUTIONS

G. A. Baker
University of California, Davis

A test of homogeneity is developed that will frequently detect nonhomogeneity by means of samples as small as fifty for some populations that are composed of normal components. The development is essentially the same as Smirnov's distribution free tests as discussed in [4].

1. INTRODUCTION

One of the most universal causes of nonnormality of populations is nonhomogeneity. This fact was pointed out by Karl Pearson [8] in 1894 who considered dissecting nonnormal populations into normal components. Random sampling from such populations was discussed by Baker [1] in 1930 with further contributions later. A test of homogeneity for normal populations was given by Baker [2] in 1941 which was based on ideas developed by Neyman [7] in 1937. The problem is still recognized and recurrent as indicated by Preston [9] and Rao [10]. For Rao it appears as a problem of classification after nonhomogeneity is demonstrated or assumed.

Over a period of time ideas of testing hypotheses concerning cumulative frequency distributions have been developed by many authors. These developments are partially indicated by Bartlett [3], Gumbel [5], Birnbaum [4], and Kac, Kiefer, and Wolfowitz [6].

These ideas are combined with empirical samplings from six hypothetical populations to develop a method for testing for nonhomogeneity and to give some idea of its effectiveness under rather widely varying population situations.

2. SAMPLED POPULATIONS

The populations from which samples were drawn are:

$$\text{(N)} \quad \frac{1296}{5\sqrt{2\pi}} \exp. - 1/2 \frac{(x - 15.5)^2}{25}$$

$$\text{(I)} \quad \frac{648}{5\sqrt{2\pi}} \left(\exp. - 1/2 \frac{(x - 15.5)^2}{25} + \exp. - 1/2 \frac{(x - 32.5)^2}{25} \right)$$

$$\text{(II)} \quad \frac{432}{5\sqrt{2\pi}} \left(\exp. - 1/2 \frac{(x - 30.5)^2}{100} + \exp. - 1/2 \frac{(x - 41.5)^2}{25} \right)$$

$$\text{(III)} \quad \frac{324}{5\sqrt{2\pi}} \left(3 \exp. - 1/2 \frac{(x - 15.5)^2}{25} + \exp. - 1/2 \frac{(x - 31.5)^2}{25} \right)$$

$$\text{(IV)} \quad \frac{432}{5\sqrt{2\pi}} \left(2 \exp. - 1/2 \frac{(x - 30.5)^2}{100} - \exp. - 1/2 \frac{(x - 35.5)^2}{25} \right)$$

$$\text{(V)} \quad \frac{648}{5\sqrt{2\pi}} \left(\exp. - 1/2 \frac{(x - 15.5)^2}{25} + 1/2 \exp. - 1/2 \frac{(x - 43.5)^2}{100} \right).$$

The method of drawing samples is the same as that originally described in [1]. Each of the populations has a total area of 1296. Thus, 1296 integers distributed over a proper range and with the frequencies indicated by the corresponding areas were entered on charts with 6 big rows and 6 big columns of squares which were subdivided into 6 little rows and 6 little columns. In each case the 1296 integers were distributed in a nonsystematic way among the 1296 individual spaces. By throwing 4 differentiated dice it was possible to draw random individuals from populations that are approximately N, I, II, III, IV, and V. The sample sizes considered are 50 and 100.

The samplings were completely independent for each population and the numbers of samples for each were not necessarily the same. The dice were ordinary dime store dice and may have shown 6's oftener than they should. Since the items were scattered and not entered in blocks, little bias seems to have appeared in the final results of the distributions of means and variances examined elsewhere.

3. TEST OF HOMOGENEITY

For a finite sample n we first order the observed values x_k from the smallest to the largest. Next, we compute the mean and standard deviation of the sample and express the observed values as deviations from the mean in terms of the standard deviation. Then, we use the tables of the normal areas to compute z_k, the area from $-\infty$ to the standardized observed values. We then consider

$$D^j = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{k}{n} - z_k\right|^j \tag{3.1}$$

or

$$D = \text{max.}\left|\frac{k}{n} - z_k\right|$$

as a test of homogeneity.

If the sample is infinite we have

$$D^j = \int_0^1 \left|\int_0^y \int_{-\infty}^x f(t)dtdx - \int_0^y \int_{-\infty}^x g(t)dtdx\right|^j dy \tag{3.2}$$

where

$$f(x) = \frac{p}{\sigma_1\sqrt{2\pi}}\,\text{exp.} - 1/2\,\frac{(x+m_1)^2}{\sigma_1^2} + \frac{q}{\sigma_2\sqrt{2\pi}}\,\text{exp.} - 1/2\,\frac{(x-m_2)^2}{\sigma_2^2}, \tag{3.3}$$

$$p + q = 1 = \mu_0$$

$$\mu_1 = -pm_1 + qm_2 = 0$$

$$\mu_2 = p\sigma_1^2 + q\sigma_2^2 + pq(m_1 + m_2)^2$$

$$m_1 > 0, \qquad m_2 > 0$$

and

$$g(x) = \frac{1}{(p\sigma_1^2 + q\sigma_2^2 + pq(m_1 + m_2)^2)^{1/2}\sqrt{2\pi}} \cdot \exp. - 1/2 \frac{x^2}{p\sigma_1^2 + q\sigma_2^2 + pq(m_1 + m_2)^2}. \tag{3.4}$$

The values of D, D^1, and $\sqrt{D^2}$ for the infinite case can be approximated by using tables for the normal distribution and repeated numerical and/or graphical integration.

4. THEORETICAL AND EMPIRICAL SAMPLING RESULTS

First attention is called to the correlation between D, D^1, and D^2 for samples of 50 as given in Table 553. It is noted that the correlation between D^1 and $\sqrt{D^2}$ is very high, averages 0.985 for the six populations. For this reason further discussions are limited to D and D^1. It is seen that D is a maximum ordinate and that D^1 is essentially an average of the absolute ordinates.

In Table 554a we compare the mean maximum ordinates D and the mean average absolute ordinates D^1 for samples of 50 and 100 and for (3.2) for the six populations.

It seems that it might be of interest to present the frequency distributions of D and D^1 in some detail and this is done in Tables 554b and 555. This presentation permits an estimate of the power of these tests to be made by inspection. To facilitate the comparisons of the two measures of nonhomogeneity for the various populations we give the per cent of the frequencies beyond 0.13 for D and beyond 2.1 for 50 D^1. We use 50 D^1 simply to avoid the bother of division of so many items by 50.

In Table 556 we give results for samples of 100 from populations N and I for D and 50 D^1.

5. DISCUSSION

We see that, even for samples of 50, for all populations D^1 and $\sqrt{D^2}$ are almost perfectly correlated. Thus we have considered only D and D^1 in detail.

TABLE 553

CORRELATIONS BETWEEN D, D^1, $\sqrt{D^2}$ FOR SIX POPULATIONS FOR SAMPLES OF 50

Population	No. of Samples	$r_{DD'}$	$r_{D\sqrt{D^2}}$	$r_{D'\sqrt{D^2}}$
N	100	.883	.918	.988
I	83	.826	.885	.987
II	83	.852	.903	.990
III	85	.896	.932	.988
IV	83	.840	.905	.983
V	83	.731	.827	.973

TABLE 554a

MEAN MAXIMUM ORDINATES, D, AND MEAN AVERAGE ABSOLUTE ORDINATES, D^1, COMPARED FOR SIX POPULATIONS FOR SAMPLES OF 50, 100, AND ∞

Population	β_1	β_2	No.	Samples of 50		Samples of 100		Values of (3.2)	
				D	D^1	D	D^1	D	D^1
N	0	3	100	.099	.032	0.086	.025	0	0
I	0	1.896	83	.128	.051	0.110	.043	.177	.072
II	.205	2.856	83	.101	.037			.485	.117
III	.378	2.712	85	.137	.050			.217	.109
IV	.123	2.989	83	.107	.036			.477	.257
V	.144	1.913	83	.168	.062			.975	.657

TABLE 554b

FREQUENCY DISTRIBUTIONS OF D FOR SAMPLES OF 50 FOR SIX PARENT POPULATIONS

Class interval	N	I	II	III	IV	V
.05–.06	3		4		3	
.06–.07	6		9		9	
.07–.08	19	2	9	3	6	
.08–.09	12	4	11	1	11	
.09–.10	17	6	14	4	9	
.10–.11	13	6	9	8	7	1
.11–.12	10	14	5	13	9	4
.12–.13	9	12	10	7	8	7
.13–.14	4	13	3	10	10	6
.14–.15	1	8	2	10	6	10
.15–.16	3	8	3	6	0	8
.16–.17	1	4	3	7	3	8
.17–.18	2	4	0	7	1	7
.18–.19		0	0	5	1	12
.19–.20		1	0	4		8
.20–.21		1	1			2
.21–.22						6
.22–.23						0
.23–.24						1
.24–.25						2
.25–.26						0
.26–.27						0
.27–.28						0
.28–.29						1
Total	100	83	83	85	83	83
Mean	.099	.128	.101	.137	.107	.168
Standard Deviation	.026	.027	.030	.031	.030	.034
Per cent > .13	11	47	14	58	25	86

Frequency distributions can be characterized, at least crudely, by their values of β_1 and β_2. For this reason β_1 and β_2 are given for the parent populations in Table 554a.

The mean values for samples of 50 and 100 and the approximate theoretical values for (3.2) given in Table 554a show that the populations vary widely in the rapidity with which maximum discrimination is attained. For instance for population I the mean value of D^1 for samples of 50 is 71 per cent of the

TABLE 555

FREQUENCY DISTRIBUTIONS OF 50 D^1 FOR SAMPLES OF 50 FOR SIX PARENT POPULATIONS

Class interval	N	I	II	III	IV	V
.9–1.0	2		1	1	4	
1.0–1.1	3		2	0	5	
1.1–1.2	5	1	7	0	5	
1.2–1.3	10	1	5	0	3	
1.3–1.4	13	0	5	1	5	
1.4–1.5	16	2	9	0	8	
1.5–1.6	8	5	6	4	5	1
1.6–1.7	12	3	5	4	6	0
1.7–1.8	8	2	3	2	6	1
1.8–1.9	4	1	6	5	4	0
1.9–2.0	6	3	5	6	6	0
2.0–2.1	2	3	2	3	4	2
2.1–2.2	2	4	10	5	3	2
2.2–2.3	2	4	2	4	3	3
2.3–2.4	3	2	2	4	4	3
2.4–2.5	2	5	3	5	2	4
2.5–2.6	0	6	1	5	2	5
2.6–2.7	1	3	1	2	0	4
2.7–2.8	0	5	0	5	1	3
2.8–2.9	0	10	0	3	3	6
2.9–3.0	0	7	1	7	2	5
3.0–3.1	0	3	1	1	1	5
3.1–3.2	1	2	1	3	0	6
3.2–3.3		3	1	3	0	4
3.3–3.4		0	1	2	0	2
3.4–3.5		2	0	3	0	5
3.5–3.6		2	2	1	0	3
3.6–3.7		1	0	2	0	8
3.7–3.8		0	0	3	0	3
3.8–3.9		1	1	0	0	4
3.9–4.0		1		1	1	3
4.0–4.1		1				1
Total	100	83	83	85	83	83
Mean	1.596	2.542	1.862	2.507	1.794	3.038
Standard Deviation	.390	.642	.633	.662	.588	.571
Per cent >2.1	11	75	33	69	27	95

value for (3.2) while for population V the mean value of D^1 for samples of 50 is only 9 per cent of the value for (3.2).

Table 555 indicates that D^1 is a better discriminator of nonhomogeneity for samples of 50 for all populations considered. However, the improvement is very slight for population IV and we cannot conclude that D^1 is better for all populations and all sample sizes. Population IV is somewhat more unusual than the other types so we might conclude that D^1 is to be generally preferred.

It is clearly indicated that, by using the techniques developed here, nonhomogeneity of sampled populations can often be detected by using samples as small as fifty.

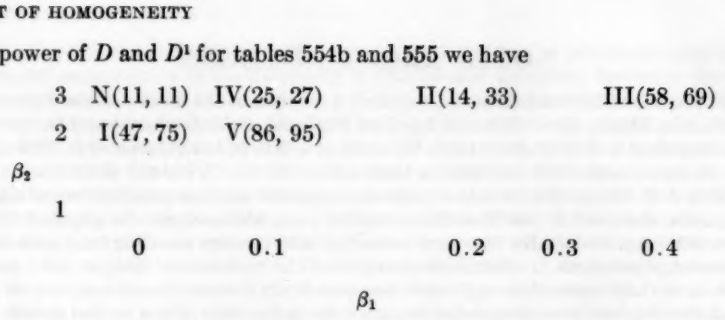
A referee points out the additional desirable discussion. The power of the test to detect nonhomogeneity depends on the alternative distribution and on the sample size. The alternatives are here characterized by β_1 and β_2. If we plot

TABLE 556

FREQUENCY DISTRIBUTIONS OF D AND $50D^1$ FOR SAMPLES OF 100 FOR A NORMAL POPULATION (N) AND POPULATION I

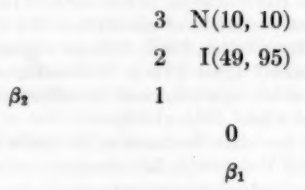
Class Interval	D		Class Interval	$50D^1$	
	N	I		N	I
0.05–0.06	3		0.8–0.9	1	
0.06–0.07	7		0.9–1.0	7	
0.07–0.08	10	1	1.0–1.1	8	
0.08–0.09	14	5	1.1–1.2	12	
0.09–0.10	3	6	1.2–1.3	5	
0.01–0.11	8	9	1.3–1.4	5	
0.11–0.12	3	9	1.4–1.5	3	2
0.12–0.13	2	5	1.5–1.6	4	0
0.13–0.14		6	1.6–1.7	3	1
0.14–0.15			1.7–1.8	0	4
			1.8–1.9	2	4
			1.9–2.0		4
			2.0–2.1		4
			2.1–2.2		7
			2.2–2.3		4
			2.3–2.4		1
			2.4–2.5		2
			2.5–2.6		3
			2.6–2.7		2
			2.7–2.8		0
			2.8–2.9		1
			2.9–3.0		0
			3.0–3.1		0
			3.1–3.2		1
			3.2–3.3		1
Total	50	41		50	41
Mean	0.086	0.110		1.249	2.166
Standard Deviation	0.018	0.016		0.259	0.411

the power of D and D^1 for tables 554b and 555 we have



where the figures in the parentheses are the per cents of samples which are >0.13 for D and >2.1 for 50 D^1. It appears that except for an irregularity for D for populations IV and II the power increases as β_1 increases and as β_2 decreases from 3.

The similar plot for samples of 100 for populations N and I is as follows:



where the figures in parentheses are the per cents of samples with $D>0.11$ and 50 $D^1>1.6$. As is to be expected samples of 100 are much more effective in detecting nonhomogeneity than are samples of 50. This is especially true for D^1 when the alternative is a population I.

We have not as yet been able to carry this investigation further.

REFERENCES

[1] Baker, G. A., "Random sampling from nonhomogeneous populations," *Metron*, Vol. 8 (1930), pp. 68–89.

[2] Baker, G. A., "Test of homogeneity for normal populations," *Annals of Mathematical Statistics*, Vol. 12 (1941), pp. 233–6.

[3] Bartlett, M. S., *An Introduction to Stochastic Processes*, Cambridge University Press, London 1955.

[4] Birnbaum, Z. W., "Distribution-free tests of fit for continuous distribution functions," *Annals of Mathematical Statistics*, Vol. 24 (1953), pp. 1–8.

[5] Gumbel, E. J., "Simple tests for given hypotheses," *Biometrika*, Vol. 32 (1942), pp. 317–33.

[6] Kac, M., Kiefer, J., and Wolfowitz, J., "On tests of normality and other tests of goodness of fit based on distance methods," *Annals of Mathematical Statistics*, Vol. 26 (1955), pp. 189–211.

[7] Neyman, J., " 'Smooth test' for goodness of fit," *Skandinavisk Aktuariedskrift* (1937) pp. 149–99.

[8] Pearson, Karl, "Contribution to the mathematical theory of evolution," *Philosophical Transactions of the Royal Society of London*, Series A, Vol. 185 (1894), pp. 71–110.

[9] Preston, Eric J., "A graphical method for the analysis of statistical distributions into two normal components," *Biometrika*, Vol. 40 (1953), pp. 460–4.

[10] Rao, C. R., *Advanced Statistical Methods in Biometric Research*, John Wiley and Sons, New York 1952.

NOTES ABOUT AUTHORS

GEORGE ALLEN BAKER, 54, has been a member of the faculty of the University of California, Davis, since 1937, and has been Professor of Mathematics and Statistician in the Experiment Station since 1955. He received a B.S. in animal science in 1926 and a Ph.D. in mathematics and statistics in 1929, both from the University of Illinois. It was there that A. R. Crathorne, his major professor, suggested work on combinations of normal populations which led to his "Random sampling from nonhomogeneous populations" in *Metron* (1930, pp. 68–89). He "has been concerned with random sampling from nonnormal and normal populations . . . distributions of means and variances of samples from combinations of normal populations and other nonnormal populations, transformations of nonnormal distributions into normal distributions, the probability that a second sample will differ from a first sample by a certain amount, correlations in nonnormal multivariate populations, linear regression, populations with means and standard deviations independently distributed, and the effect of selection on linear functions of normally correlated variables." Before going to Davis, Baker was a Fellow at Columbia University 1929–30 and later taught mathematics at Shurtleff College and Mississippi Woman's College and served as statistician with the New York State Department of Health and the United States Department of Agriculture. His numerous papers include four in the *JASA:* "The significance of the product-moment coefficient of correlation with special reference to the character of the marginal distributions" (Dec. 1930), "Linear regression when the standard deviations of arrays are not constant" (Dec. 1941), "Correlations between functions of variables" (Dec. 1942), and "Test of the significance of the differences of per cents of emergence of seedlings in multiple field trials" (Mar. 1945).

ISADORE BLUMEN, 38, is Associate Professor of Statistics in the School of Industrial and Labor Relations at Cornell University. Between graduation from the University of Minnesota in 1941 and post-war return to graduate study, he spent about two years as a statistician for the Federal government (principally with the War Production Board), with the remaining time in military service. He received an M.A. and a Ph.D. in mathematical statistics at the University of North Carolina in 1948 and 1954. His "On the ranking problem" appeared in *Psychometrika* (Mar. 1957) and he is co-author of a monograph, *The Industrial Mobility of Labor as a Probability Process.* He has published previously in the *JASA* as a reviewer. His interests and experience have been in applications of statistics in the social sciences. The present paper is an outgrowth of interests in nonparametric statistics and multivariate analysis.

JAMES VANDIVER BRADLEY, 34, is Research Psychologist, Aero Medical Laboratory, Wright Air Development Center, Ohio. After work in electrical engineering at Lehigh in 1943–44, he majored in physics for a B.A. at Washington and Lee in 1947, took an M.A. in experimental psychology at the University of Virginia in 1949, and had later work in psychology at Wisconsin and Ohio. His fields of interest are experimental psychology, distribution-free statistics, and probability. His present work is research on how best to adapt the mode by which an instrument is operated to the human capabilities of a general class of operators. His "Control knob arrangement can save aircraft instrument panel space" appeared in *Journal of Aviation Medicine* (1957, pp. 322–7). The present article, his first in the *JASA*, followed use of a trial and error technique for balancing out immediate sequential effects in Latin square designs. With increasing square size the work became increasingly time consuming and frustrating, so the method described was worked out, and is presented for other research workers who may be losing time by trial and error methods. The sequential effects with which he was concerned were primarily learning and motivation, both of which are influenced not only by the position in sequence of an experimental condition but also by the nature of the immediately preceding condition. He attempts to balance out contrary effects on subjects in order not to bias his experiments.

ACHESON JOHNSTON DUNCAN, 53, is Associate Professor of Statistics at The Johns Hopkins University, with experience in statistics from undergraduate days. He has a B.S. in mathematics (1925) and an M.A. (1927) and a Ph.D. (1936) in economics, all from Princeton. From 1929 to 1942 and from 1945 to 1946 he was a member of the Eco-

nomics Department at Princeton with one year (1931–32) out for advanced study in statistics and econometrics at the University of Chicago and Columbia. Service in the U. S. Army occupied the years 1942–45. He is a Fellow of the American Statistical Association and of the American Society for Quality Control, a member of the Institute of Mathematical Statistics and the Econometric Society, and serves on the editorial board of *Industrial Quality Control*. His published books are *Elementary Statistics and Applications* and *Sampling Statistics and Applications* (co-author), McGraw-Hill, 1944 and 1945; and *Quality Control and Industrial Statistics*, Richard D. Irwin, 1952. His previous publications in the *JASA* are: "The economic design of $\bar{X}$ charts used to maintain current control of a process" (June, 1956) and "Charts of the 10% and 50% points of the operating characteristic curves for fixed effects analysis of variance F tests, $\alpha = 0.01$ and 0.05." (Sept. 1957). The present article is the result of his attempt to write up the work of Lieberman and Resnikoff to present in concise form the steps necessary to enable an individual to design a double limit variables sampling plan to meet criteria of his own choosing.

MARILYN DUNSING, 32, is Assistant Professor of Home Economics, University of California, Davis. She received the M.B.A. in 1948 and the Ph.D. in home economics in 1954 from the University of Chicago. While working for her degree, she taught in a junior college of Chicago. This first appearance in the *JASA* is an outgrowth of work on her thesis.

OTTO DYKSTRA, JR., officially an "Associate Technologist" at General Foods Research Center in Tarrytown, New York, has been doing statistical work for General Foods since 1954 when he returned to civilian life from the Analytical Section of the Development and Proof Service, Aberdeen Proving Grounds. He is now 28. He received a B.A. from Columbia in mathematics in 1951 and an M.S. in applied statistics from Rutgers in 1957. His fields of special interest are experimental design and taste testing. He has not published previously in the *JASA*, but has published "A note on the rank analysis of incomplete block designs—applications beyond the scope of existing tables" in *Biometrics* (Sept. 1956) and "Applications of Scheffé's method of paired comparisons in food quality evaluation" in the *Transactions* of the 1957 Convention of the American Society for Quality Control.

LILA ELVEBACK, 42, is Associate Professor of Biostatistics in the Department of Tropical Medicine and Public Health, Tulane University. She received a B.A. in mathematics at Minnesota in 1941, an A.M. in mathematical statistics at Columbia in 1948, and a Ph.D. in statistics at Minnesota in 1955. She has taught mathematics at Minnesota and biostatistics at Columbia and Minnesota. During the war she worked in the Optical Instruments Section of the Office of Scientific Research and Development. Her present position dates from 1955. Articles by her have appeared in medical journals, but this is her first in the *JASA*. She became acquainted with this general problem, which was suggested by Joseph Berkson, during the summer of 1951 while working on a model for relapse-recovery studies in chronic disease under the direction of Jerzy Neyman. During the following year she spent six months on fellowship in the Statistics Section of the Mayo Clinic.

HARRY ERNST, 36, is Director of Research of the Institute for Economic Research of Tufts University. He received a Ph.D. from Harvard in 1955. In 1954 he became a consultant to the Federal Reserve System, and in 1956 an Assistant Professor and in 1958 an Associate Professor at Tufts. His interests are (1) measurement and prediction of productivity and production and (2) plant efficiency and income distribution systems. His "Accounting for productivity changes" appeared in the *Harvard Business Review* (May–June 1956), and his "Functional analysis, plant productivity and efficiency" in *Productivity Measurement Review* (May 1957). The present article, his first in the *JASA*, grew from his surprise at finding no detailed current data available in the U. S. on kilowatt-hours, which he has used as indicators of production changes.

ROBERT FERBER, 36, received a B.S. at City College of New York in 1942 and an M.A. and Ph.D. at the University of Chicago in 1945 and 1951, with graduate work at Columbia in mathematical statistics in 1946–47. His fields of interest are statistics, economics, and marketing. He is now Professor of Economics in the Department of Economics

and Research Professor of Economics in the Bureau of Economic and Business Research at the University of Illinois. He is also director of the Ford Foundation-financed Consumer Savings Project sponsored by the Inter-University Committee for Research on Consumer Behavior. Before going to Urbana in 1948, he was research assistant for the Cowles Commission, chief statistician for Industrial Surveys Company, and economist and statistician for I. de Vegh in New York. He has published *Statistical Techniques in Market Research*, McGraw-Hill, 1949, *Marketing Research—Selected Literature* (co-editor), Wm. C. Brown and Co., 1952, *Cases and Problems for Marketing Research* (co-author), Ronald Press, 1954, *A Basic Bibliography on Marketing Research* (co-author), American Marketing Association, 1956, *Motivation and Market Behavior* (co-author), and has in preparation *Research Methods in Economics and Marketing* (co-author), and *Economics in Modern Life* (co-author). In the *JASA* he has published "Measuring the accuracy and structure of businessmen's expectations" (June 1953), "On the reliability of responses secured in sample surveys" (Sept. 1955), and "The effect of respondent ignorance on survey results" (Dec. 1956). "The present article is part of a larger study on the Illinois employers labor force anticipations which was initiated as part of the project on 'Expectations and business functions' carried out at the University of Illinois some years ago under the direction of Franco Modigliani, with financial help from the Merrill Foundation."

LEO A. GOODMAN, 29, has been Professor of Statistics and Sociology at the University of Chicago since 1955. He began teaching at Chicago in 1950 after receiving the M.A. and Ph.D. in mathematics and mathematical statistics at Princeton the same year. His A.B. in mathematics and sociology was received at Syracuse in 1948. In 1953–54 he was at Cambridge University on a Fulbright Award. He is a Fellow of the Institute of Mathematical Statistics and of the American Statistical Association. His numerous publications include the following in the *JASA:* "Serial numbers analysis" (Dec. 1952), "Methods of measuring useful life of equipment under operational conditions" (Sept. 1953), "Some practical techniques in serial number analysis" (Mar. 1954), and "Measures of association for cross classifications," (co-author) (Dec. 1954).

HAROLD DONALD GRIFFIN, 65, is "free-lance," which he explains means retired, with a permanent address of Box 1411, Lincoln 1, Nebraska. He received an A.B. at Bethany College in 1919, and then had varied teaching and administrative assignments at Missouri and North Dakota colleges before receiving a Ph.D. at the University of Missouri in 1929. His professional interests are indicated by the title of his doctoral dissertation, "Charts and Tables for Prediction in Psychology and Education." He now spends much of his time in travel (with study of the environment) and photography, particularly stereophotography. He has taken colored stereo pictures since 1924. Included in his list of publications are two previous articles in the *JASA:* "Nomogram for correcting simple and multiple correlation coefficients" (Sept. 1930), and "On the coefficient of part correlation" (Sept. 1932). His explanation of interest in the subject of his present article is:

> "When Peter Sandiford's 'Educational Psychology' came to my attention during the winter of 1928–29, I saw possibilities in Sidney D. Holmes' formula for graphic '*r*' as a measure of rank correlation with as much validity for that purpose as the conventional Spearman method. Although Dr. Sandiford, with whom I corresponded, could not see it that way, I never have ceased to use the Holmes' formula in my own private studies. I was not familiar with Kendall's studies of rank correlation methods until I read Harold F. Bright's 'A method for computing the Kendall tau coefficient,' in the winter number of *Educational and Psychological Measurement* in 1954. I saw at once that my favorite rank correlation method had scholarly support at last. However, when I read M. G. Kendall's book, I could not help but feel that Holmes' graphic method, at least for modest numbers of ranks, was simpler to compute than any of Kendall's published methods."

During World War II he was a clinical psychologist in the Army, and from 1945 to 1956 he worked for the Veterans Administration in Nebraska.

HERMAN EZRA GUTERMAN, 38, is a mathematical statistician with the Internal Revenue Service. His address is 43 Gallatin Street N.W., Washington, D. C. He received a B.A. in mathematics from the University of Scranton in 1942, and did graduate work at

the University of Pittsburgh and at George Washington University in mathematical statistics. His principal interest is in sampling.

HERMAN OTTO HARTLEY, 46, is Professor in the Department of Statistics and Statistical Laboratory at Iowa State College, Ames. He received a Ph.D. in mathematics at the University of Berlin 1934, a Ph.D. in mathematical statistics at the University of Cambridge 1940, and a D.Sc. in mathematical statistics at the University of London 1954. He had experience in England as a college statistician, as scientific officer for a firm of scientific consultants engaged in organizing large-scale computations for the British Army and Navy and the American Air Force during World War II, and as a lecturer in statistics at University College, while also acting as statistical consultant to the Royal Cancer Hospital. He went to Ames on leave as a Visiting Professor in 1953 and became Professor of Statistics there in 1954. He is a Fellow of the Institute of Mathematical Statistics and of the American Statistical Association, and a member of International Statistical Institute. He is a co-author of *Biometrika Tables for Statisticians*, Cambridge University Press 1954, and is a frequent contributor to *Biometrika*, *Biometrics*, and the *Annals of Mathematical Statistics*. The substance of the first half of the present paper was presented at a regional meeting of the Institute of Mathematical Statistics. The result was a discussion with Leo Goodman, leading to this joint paper. His special interests are sampling problems and high speed computing.

IRENE HESS, 47, is Assistant Head of Sampling Section, Survey Research Center, University of Michigan. She took an A.B. at Indiana in mathematics in 1931, and has since studied statistics and mathematics at Michigan and George Washington. She taught high-school mathematics before going to Washington, D.C., to work with the U. S. Geological Survey during the War. She spent one year at the National Bureau of Standards before going into the Sampling Section, Population Division at the Bureau of the Census. Since 1954 she has been at Ann Arbor. Her fields of interest are in the "design and planning of sample surveys, the sample selection, methods of improving and controlling the sampling activities of interviewers, methods of estimation, computation of sampling variability, research directed toward improving the efficiency of sample surveys."

JONATHAN ROBERTS TYSON HUGHES, 30, is Associate Professor of Economics at Purdue University and this year has been Visiting Assistant Professor of Economics at Columbia. After receiving the B.S. at Utah State University, he did graduate work 1950–52 at the University of Washington and then spent two and one-half years at Oxford as a Rhodes Scholar, taking a D.Phil. He was an economist in the research department of the Federal Reserve Bank of New York 1955–56. He has had published "The commercial crisis of 1857," *Oxford Economic Papers*, June 1956, and "The Canadian dollar premium," *Business Scope*, May 18, 1957. The main statistical source for the present article was encountered by accident at Oxford in the fall of 1954. It suggested the writing of a more definitive account than was available of British steam navigation before 1860, but there were many problems, including the immense one of reordering the data. Nothing was done about it until Stanley Reiter at Purdue suggested a kind of statistical analysis that would make the data more meaningful. Thus began their collaboration. When the authors began their computations, using punch-cards and IBM machines, on almost every sorting the machine would initially "throw out" a single card, that for the "Great Eastern." "And so a century after she was built . . . the Great Iron Ship was still too large, too cumbersome, unmanageable . . . in her only remaining form—as a set of statistics. In the end we had to make all sorts of allowances for the (relatively) huge dimensions of the ship." The author's fields of interest are: economic history, monetary economics, economic development, and business fluctuations.

HOWARD LEWIS JONES, 60, is general supervisor of statistics with the Illinois Bell Telephone Company, having started there as a clerk in 1923. He received a B.A. at the University of Illinois in 1923, has studied and been a part-time instructor in statistics at the University of Chicago. He was an Associate Editor of the *JASA* 1951–53; is a Vice-President and a Fellow of the American Statistical Association and a Fellow of the American Society for Quality Control; has recently been appointed to the editorial board of *Industrial Quality Control;* and has served on many important committees. In the *JASA* his articles are: "The use of grouped measurements" (Dec. 1941), "Fitting polynomial

trends to seasonal data by the method of least squares" (Dec. 1943), "Linear regression functions with neglected variables" (Sept. 1946), "Approximating the mode from weighted sample values" (Mar. 1953), "Application of sampling procedures to business operations" (Sept. 1955), and "Investigating the properties of a sample mean by employing random subsample means" (Mar. 1956). About his present article he makes three points:

1. It may make statisticians aware that we have a real problem in deciding what to do about unsatisfactory samples, and that we are on tenuous ground if we simply throw away a sample we don't like and select another.

2. It may cause statisticians to give greater preference to sample designs where the probability of selecting an inadmissible sample is small.

3. It may stimulate further research that will lead to a better solution to the problem.

EDWARD LYNN KAPLAN, 38, is a mathematician in the Computation Division of the University of California Radiation Laboratory at Livermore, where he has been interested in Monte Carlo methods suitable for neutronics problems. He graduated in mathematics from the Carnegie Institute of Technology in 1941, and until 1948 was a mathematician at the Naval Ordnance Laboratory. In 1950 he obtained a Ph.D. in mathematical statistics at Princeton and then became a member of the Mathematical Research Department of Bell Telephone Laboratories, where he was concerned with the analysis of military systems and statistical techniques. He has published several articles, including four in the *Journal of Mathematics and Physics*. The collaboration of the present article resulted when Kaplan and Meier almost simultaneously prepared separate manuscripts that were in many respects similar.

BERNARD DOV KARPINOS, 54, is Consultant on induction activities to the Chief of the Medical Statistics Division of the Surgeon General's Office, Army. He is responsible for the processing and analysis of the individual physical examination records of men processed for military service in the Armed Forces. He is also Chief of the Prevalence Studies Section. During World War II he was inducted in the Army as a private, but upon his release was a Major. He has also been connected with the U. S. Public Health Service, the Illinois Tax Commission, the National Resources Committee, the Iowa State Planning Board, and the Iowa Bureau of Business Research. After a childhood in the Ukraine, Karpinos came to the United States alone in his early twenties with no money and no knowledge of English, contacted a friend from Russia, and worked for a while on a Hebrew weekly paper. He left for Iowa City after hearing of a Hebrew teaching position there, and entered the State University as a freshman in 1928. He earned a bachelor's degree in 1930, an M.A. in 1932, and a Ph.D. in 1934, with economics and sociology as major subjects. He has a long list of published studies, indicating his shift in interests from population problems (descriptive and theoretical) to general health problems, and then to military problems. This is his first publication in the *JASA*.

LESLIE KISH, 47, is research statistician and Head of the Sampling Section, Survey Research Center, and Associate Professor of Sociology at the University of Michigan. He received a B.S. in 1939 at the College of the City of New York in mathematics, after having studied physics and engineering. For several years he worked on statistical problems in biological research. After war service as a meteorologist, he studied statistics and received an M.A. in mathematical statistics in 1948 and a Ph.D. in sociology in 1952 at the University of Michigan. His interests include sample design, social statistics, and population studies. The *JASA* has published his "Response error in estimating the value of homes" (co-author, Sept. 1954), "Controlled selection—a technique in probability sampling (co-author, Sept. 1950), and "A procedure for objective respondent selection within the household" (Sept. 1949). The present article was started about three years ago to document a serious and previously little known source of error in samples. The authors were able to put a happy end to their work because "as a result of our search for the sources of error, we were able to reduce the errors from large to relatively small."

ALFRED JEROME McKEON, 31, is a mathematical statistician in the Statistics Division of the Internal Revenue Service, Washington 25, D. C. He has an A.B. from George Washington University, 1949, and has taken further work there in mathematics and statistics. He began statistical work with the Government in 1950. He is co-author of

"A comparison of serial number digit sampling with systematic and random sampling," *Applied Statistics* (Mar. 1953).

PAUL MEIER, 33, is Associate Professor of Statistics at the University of Chicago. He received a B.A. at Oberlin College in physics in 1945, and a Ph.D. at Princeton in mathematics in 1951. He was Assistant Professor of Mathematics at Lehigh University 1948–49, and was with the Statistical Research Department of the Philadelphia Tuberculosis and Health Association 1949–51. He was Research Associate at Princeton 1951–52, and at Johns Hopkins 1952–57, was first Research Associate, then Assistant Professor and Associate Professor of Biostatistics. He lists his interests as theoretical statistics, estimation and experimental design, and applications of statistics, particularly to biology and medicine. He published "Analysis of simple lattice designs with unequal sets of replications" in the *JASA* (Dec. 1954).

GEOFFREY HOYT MOORE, 44, has been Associate Director of Research, National Bureau of Economic Research, since 1948, starting as a member of the research staff in 1942. He received a B.S. and an M.S. from Rutgers in 1933 and 1937, specializing in farm credit, and a Ph.D. from Harvard in 1947 with a thesis on harvest cycles. He was an instructor at Rutgers from 1936 to 1942, an Associate Professor of Economics at New York University 1947–48, and a Lecturer at Columbia 1953–54. Since 1953 he has been a consultant to the Council of Economic Advisers, Washington, D. C. He is author of National Bureau Occasional Paper 18, *Production of Industrial Materials in World Wars I and II*, and Occasional Paper 31, *Statistical Indicators of Cyclical Revivals and Recessions*, both issued in 1950. He has published in the *JASA*: "A significance test for time series analysis" (co-author, Sept. 1941), "Time series significance tests based on signs of differences" (co-author, June 1943), and "Industrial materials production, World Wars I and II" (Sept. 1944). His interests are business cycles, forecasting, and income distribution. About the present paper he writes:

"Something of the history of this project may be of interest since it has survived two (or almost two) recessions and one intervening expansion. It began in January 1954, when I sent Arthur F. Burns, then Chairman of the Council of Economic Advisers, a table showing the changes in a dozen indicators during the first three, six, and nine months after the business cycle peaks of June 1929, May 1937, November 1948, and July 1953 (to November or December only). I pointed out that the character of a recession did not emerge within three months after it began, so far as rates of change in most time series are concerned, but it did begin to appear after six months, especially in 'leading' series, and is still clearer after nine months. Further, I noted that the 1953 recession then appeared distinctly milder than 1937 or 1929, judging by the leading series, and that it closely resembled the 1948 recession. Burns asked if we could extend the table to cover some 60 different economic series and put it on a month by month basis so that it could be kept up to date, and this was done.

"In November 1954 we constructed a similar set of tables in order to compare the business expansion that began in August with earlier business expansions. Here we found it illuminating to array the expansions according to the magnitude or depth of the preceding contractions, since expansions following severe contractions typically proceed at a more rapid pace than those that follow mild contractions. The 1954–55 expansion was placed, therefore, at the moderate end of the scale because the preceding contraction was one of the mildest in our business cycle chronology. And the pace of the 1954–55 recovery in aggregate activity was moderate, judged by historical standards, despite the fact that many observers were surprised by its vigor.

"The project was resumed in October 1957 when R. J. Saulnier, present Chairman of the Council, requested the assistance of the National Bureau in constructing tables similar to those used in 1954. Tables covering some seventy monthly and quarterly economic series were promptly prepared, and they have been kept up to date. At the same time, Julius Shiskin of the Bureau of the Census, began work for the Council along similar lines (see note 6 in my paper). Shiskin had spent the preceding year as a Research Associate of the National Bureau developing, testing, and improving electronic computer

programs to analyze current economic statistics, and he promptly put them to use in analyzing the scope and magnitude of the current recession."

MARGARET GILPIN REID, 62, is Professor in the Department of Economics of the University of Chicago. Before going to Chicago in 1951, her affiliations were with the University of Illinois, the U. S. Department of Agriculture, and Iowa State College. Her early student days were spent in Canada; she received the Ph.D. from the University of Chicago in 1931. She is author of *Economics of Household Production* (Wiley 1932), *Consumers and the Market* (Crofts, now Appleton-Century, 1938, with revised editions in 1939 and 1942), and *Food for People* (Wiley 1942). She wrote "Distribution of Income and Consumption" in *American Income and Its Uses* of the Series on Ethics and Economics of Society (Harper's 1954) and an article, "Food, liquor, and tobacco" in *American Needs and Resources* (20th Century Fund 1955). In 1943 she was on a Committee on Cost of Living for the American Statistical Association. This is her first publication in the *JASA*. Her special interest is family and consumption economics.

STANLEY REITER, 33, received an A.B. in economics at Queens College (1947), an M.A. (1950) and a Ph.D. (1955) in the same subject at the University of Chicago, and is now Associate Professor at Purdue in the Departments of Economics, Industrial Management and Transportation, and Mathematics. He was at the University of Chicago as Research Associate from 1949 to 1950, then was instructor in economics at Stanford, 1950–53. The next year he was a research associate in the Applied Mathematics and Statistics Laboratory at Stanford, then went to Purdue. He was co-author of "A model of transportation" and "On the choice of a crop rotation plan" in *Activity Analysis of Production and Allocation* (Wiley 1951). The *JASA* published his "Estimates of bounded relative error for the ratio of variance of normal distributions" (Sept. 1956), and economics journals have published other articles of his. Discussions about the present paper between the authors "kept raising new possibilities of meaning in the data and . . . more ideas about the relationships of these data to the history of the period . . . despite the impersonal nature of IBM data processing both [authors] found" they had special feelings about the "Great Eastern." "She was a wonder."

ARLYN CUSTER ROSANDER, 54, specialized in engineering, mathematics, and statistics at the Universities of Michigan and Wisconsin, receiving a B.S. in 1925 at Michigan, an M.A. in 1928 at Wisconsin, and a Ph.D. in 1933 at Chicago. Prior to 1940 he was technical director of market, consumer, and opinion surveys on a local, state, and nationwide basis. From 1940 to 1945 he applied sampling, quality control, and other statistical methods to production, materials, and manpower control in the War Production Board and predecessor organizations. Since 1945 he has been Chief, Sampling and Estimating Methods Section, Statistics Division, Internal Revenue Service, developing, applying, and directing the application of sampling, quality control, and other statistical and mathematical techniques to problems of tax management. From 1946 to 1952 he was Lecturer in statistics at George Washington University. Besides about 40 papers and booklets on applied probability statistics, including sampling, quality control, work sampling, and linear programming, he has published *Elementary Principles of Statistics* (Van Nostrand, 1951). His previous publications in the *JASA* are "The use of inversions as a test of random order" (Sept. 1942), and "Sampling 1949 corporation income tax returns" (co-author, June 1951).

MORRIS JAMES SLONIM, 49, has been Assistant for Mathematical Statistics, Directorate of Statistical Services, Headquarters, United States Air Force, since 1954. Previous to that he had been on active military duty with the Air Force in Florida, New Guinea, the Philippines, Okinawa, Korea, and Japan. Before joining the Air Force he was with the Department of Public Assistance of the State of Pennsylvania for three years. One year was spent with the Bureau of Labor Statistics. He is interested in sampling and electronic computer applications. His only previous article in the *JASA* was "Sampling in a nutshell" (June 1957). He hopes that his experience with writing for the *JASA* will stimulate other readers to review their files for material suitable for publication. He has an A.B. (1930) in mathematics from Harvard, an M.B.A. (1932) in statistics from the Harvard School of Business Administration, and is presently taking courses in statistics at the American University in Washington, D. C.

SUMMARIES OF PAPERS DELIVERED AT THE 117TH ANNUAL MEETING OF THE AMERICAN STATISTICAL ASSOCIATION IN ATLANTIC CITY, SEPTEMBER 10–13, 1957

Edited by ARMEN A. ALCHIAN, *University of California, Los Angeles*

The present section contains all available abstracts presented at the 1957 meeting of the American Statistical Association in Atlantic City. The sequence of presentation is alphabetical according to author. Several items are marked with an asterisk indicating that they are published instead in *Biometrics* for March, 1958, pages 131–144.

INDEX

ADELMAN, M. A., *A Current Appraisal of Concentration Statistics* 568
*ALEXANDER, CHESTER, *Longevity as a Family Trait* . . . . . *Bio*
ALEXANDER, SYDNEY S., *Rate of Change Approaches to Forecasting—Diffusion Indexes and First Differences*. 568
ANDERSON, R. L., *Some Aspects of Factorial Experimentation* 569
ANDREW, J. E., *Measuring Labour Stability and its Application to Employee Recruitment* 569
*ARMITAGE, PETER, *Sequential Procedures for Medical Trials* . . . . . *Bio*
ATKINSON, JAY L., *The Implications of the Highway Program as an Important Factor in the Long-Run Economic Outlook* 569
*ATWOOD, EARL L., *A Procedure for Removing the Effect of Response Bias Errors from Waterfowl Hunter Questionnaire Responses*. . . . . *Bio*
BACON, RALPH HOYT, *Probability of Runs in Correlated Data*. 569
BANCROFT, T. A., *The Demand and Supply of Statisticians from the Standpoint of Colleges*. 570
BANKS, J. E., *Selected Stock Market Indicators and Their Current Implications* . 570
BORTS, GEORGE H., *Summary of Regional Aspects of Cyclical Fluctuations* . . . 570
BROWN, MURRAY, AND HERMAN ROSEMAN, *Factors Affecting the Level of Fixed Investment in Manufacturing Industries for the Period, 1951–1955* 570
BUND, HENRY, *How an Economic Consultant Prepares Presentations for Clients* . 571
BURLEY, ORIN E., *Inventory Expectations Data*. 571
BURDOCK, E. I., *Classification by Configuration and Level* 571
*CHIANG, C. L., *An Application of Stochastic Processes to the Life Tables and the Standard Error of Age Adjusted Rates* . . . . . *Bio*
CLARK, EUGENE F., *Truncation to Meet Requirements on Means* 572
COHEN, MORRIS, *Anticipations Data in the Capital Goods Field* 572
COLM, GERHARD, *Long Range Economic Outlook* 572
CONKLIN, MAXWELL R., AND OWEN C. GRETTON, *Some Experience with Electronic Computers in Processing Economic Census Data* 572
COOMBS, CLYDE H., *Inconsistency of Preferences as a Measure of Psychological Distance* 573
COWAN, DAVID, *Modifications of the Correlation Technique for Use in a Process Capability Analysis* 573
CREAMER, DANIEL, *Methods of Inquiry—The Income Approach* 573
CROCKETT, *A New Type of Estimate of the Income Elasticity of the Demand for Food* 573
CROWDER, EDWARD T., *Accounting Records: A Limiting Factor in Economic Data Collection* 574

DALY, REX F., *Demand for Farm Products at Retail and the Farm Level* . . . 574
DANIEL, CUTHBERT, *Factorial Experiments* 574
DAVID, HERBERT T., *A Three-sample Kolmogoroff-Smirnoff Test* 574
DE BAUM, ROBERT M., *Experiences with Response Surface Experimentation* . . 575
DORSEY, HAROLD B., *Tomorrow's Investments* 575
EDMONDS, STAHRL, *Communicating Statistical Findings to Management* 575
EISEMANN, DORIS M., *Bank Credit and Inventory Cycles* 575
EISENHART, CHURCHILL, *Karl Pearson: An Appreciation of One Aspect of His Work* 576
ELLIS, IRA T., *Presentation of Statistics to the Management of a Large Chemical Company* 576
*FEDERER, W. T., *Augmented Designs* . . . . . *Bio*
*FIX, EVELYN, *Application of a Stochastic Model of Contagion to Data from a Foundling Home* . . . . . *Bio*
FLAHERTY, WILLIAM C., *Methods of Estimating Inventories of Passenger Cars* . . 576
FOOTE, RICHARD, *Avoiding Nonsense Correlations in Time Series* 576
FOSS, MURRAY F., *Interrelationships between Investment and Sales Anticipations* . 577
*FULLER, JOHN L., *Quantitative Genetics of Behavioral Attributes* 577
FUTRANSKY, DAVID L., *Supply and Demand for Statisticians in Federal Service* . . 577
GARFIELD, FRANK R., *How Results of Statistical and Economic Research are Presented to the Board of Governors of the Federal Reserve System* 577
GIVENS, MEREDITH B., *The Low Income Problem: Basic Concepts* 577
*GREENHOUSE, S. W., *Trials of Tranquilizers* . . . . . *Bio*
GREENHOUSE, SAMUEL W., AND SEYMOUR GEISSER, *Application of Analysis of Variance Techniques to Profile Analysis* 578
GREENOUGH, M. L., *Mass Data Input to Computers, a Progress Report on FOSDIC* . 578
GULLIKSEN, HAROLD, *Distinctive Characteristics of Psychometric Statistics. Problems of Metric* 578
HADER, R. J., *Statistics for Engineers at N. C. State College* 578
HAMBURG, MORRIS, *Some Experiments with Demand Relationships for Clothing* . 578
HANSON, R. H., AND E. S. MARKS, *The Influence of the Interviewer on the Accuracy of Survey Results* 579
*HANSON, W. D., *Theoretical Distributions of Undisturbed Chromosome Lengths in the Gamets of F_1 and Backcross Generations* . . . . . *Bio*
HARSHBARGER, BOYD, *The Demand and Supply of Statisticians from a Survey of Industries and Colleges* 579
*HARTER, H. L., AND MARY D. LUM, *A Note on Tukey's One Degree of Freedom for Non-additivity* . . . . . *Bio*
*HARRIS, EUGENE K., *On the Probability of Survival of Bacteria in Sea Water* . . *Bio*
HARTLEY, H. O., *Maximum Likelihood Estimation from Sampling Data* 579
*HARVEY, W. R., *Estimation of Genetic Progress in Each of n Traits* . . . . . *Bio*
*HEMPHILL, R. M., *Retrolental Fibroplasia, Clinical Trial* . . . . . *Bio*
HERMALIN, ALBERT I., *The 1955 Survey of Life Insurance Ownership* 579
HESS, IRENE, *On the Problems of Coverage in Samples of Dwelling Units* . . . 580
HICKMAN, BERT G., *An Experiment with Weighted Indexes of Cyclical Diffusion* . 580
*HOFFMAN, JOSEPH G., *The Role of Variable Generation Time in Tissue Cell Multiplication* . . . . . *Bio*
*HUDDLESTON, HAROLD F., *Use of Plant Characteristics in Objective Yield Forecasting* . . . . . *Bio*
JAHN, JULIUS A., *The Design of an Experiment to Measure the Effectiveness and Costs of Programs of Services for Aging Persons* 580
*KEMPTHORNE, O., *The Estimation of Environmental and Genetic Trends from Selected Data* . . . . . *Bio*
KNOWLES, JAMES W., *An Appraisal of Productivity Projections* 580
KOFFSKY, NATHAN M., *Farm Income Estimates—Methods and Meaning* 581
*KUZNETS, GEORGE M., AND YAIR MUNDLAK, *Objective Forecasts of Fruit Production* . *Bio*

LAMALE, HELEN HUMES, *The Contribution of Cross-Sectional Survey Data to Prediction of Aggregative Components of Consumption* . . . 581
*LANDAHL, H. D., *Some Theoretical Considerations on Potentiation in Drug Interaction* . . . *Bio*
LANDAU, EMANUEL, *Median Age at First Marriage in the United States* . . . 581
LAURENT, ANDRE G., *A Model for Failure Data* . . . 581
LEABO, DICK A., *Use of Sample Survey Data and Analysis of Variance as a Guide to Packaging and Pricing Policies* . . . 581
LEVINE, EUGENE, AND BENJAMIN BUCHBINDER, *Evaluating Influences on Satisfaction with Patient Care in Hospitals through Multiple Regression Techniques* . 582
LEVINE, LOUIS, *The Role of Industrial Classification in the Employment Security Program* . . . 582
LEW, EDWARD A., *Variations in Mortality from Heart Disease* . . . 582
LIEBERMAN, ALFRED, *Table for the Determination of Two-Sided Tolerance Limits for the Normal Distribution* . . . 582
LININGER, CHARLES, EVA MUELLER, AND HANS WYSS, *The Contribution of Panel Studies to Forecasting the Automobile Market* . . . 583
LIPPITT, VERNON G., *Research Methods for Analyzing Short-Term Demand for Consumer Durables* . . . 583
LIPSTEIN, BENJAMIN, *Some Statistical Problems in Urban Renewal* . . . 583
LOCKS, MITCHELL O., *Automatic Programming as a Means of Enabling Statistical Applications of Computers* . . . 583
LORD, FREDERIC M., *Distinctive Characteristics of Psychometric Statistics Problems Arising from Errors of Measurement* . . . 584
MANN, GEORGE V., *The Evidence Relating Diet to Coronary Heart Disease* . . . 584
MARKS, ELI S., AND BENJAMIN J. TEPPING, *Some Probability Considerations in the Estimation of Computer Running Time* . . . 584
MARSHALL, CARL E., *Greenhouse and Laboratory Experiments* . . . 584
MORRIS, FRANK E., *The Predictive Value of the National Bureau's Leading Indicators* . . . 585
MASON, DAVID D., *Field Experiments* . . . 585
MOSS, MILTON, *Industrial Activity and Productivity* . . . 585
MUELLER, EVA, *Sample Surveys as a Tool for Life Insurance Research* . . . 586
MUELLER, PAUL, *Evaluating Case History Data at Time of Admission to a Juvenile Training Institution* . . . 586
NATRELLA, VITO, *Implications of Pension Fund Accumulations* . . . 586
NEWCOMB, ROBINSON, *Summary on Housing* . . . 586
NEWELL, STERLING R., *A Program for Development of Crop and Livestock Estimates* . 587
ORWICZ, LEO M., *Labor Market Experience after Exhaustion of Unemployment Benefits* . . . 587
*PARR, ROBERT, AND LYLE CALVIN, *Forecasting Filbert Production from Counts of Nuts Set and Weights of Nuts* . . . *Bio*
PENDERGRASS, ROBERT NIXON, *Ranking in Triple Comparisons* . . . 587
PLUNKETT, MARGARET L., *The School Leaver in the World of Work* . . . 587
POWELL, H. R., *Methods of Estimating Retail and Consumer Inventories* . . . 588
RILEY, H. E., *Seasonality in the Consumer Price Index* . . . 588
*ROBSON, D. S., *Applications of Multivariate Polykays to Genetic Covariance Component Analysis* . . . *Bio*
ROTTENBERG, IRVING, *New Statistics on Companies and on Concentration in Manufacturing from the 1954 Census* . . . 588
SAUNDERS, SAM C., *Sequential Distribution-free Tolerance Regions* . . . 588
SAVILLE, LLOYD, *A Price Index for Custom Made Products* . . . 589
SCHAAF, ALBERT H., *APHA Housing Surveys and Rehabilitation Costs* . . . 589
SCOTT, IRA O., *The Implications of the Changing Ownership of Federal Securities* . 589
SEGEL, DAVID, *Drop-outs in High School in Large Cities* . . . 589

SELEGAN, GALENA V., AND ROOF, MICHAEL K., *Russian Population Censuses: Methods and Concepts* 590
SHISKIN, JULIUS, *Seasonal Adjustments of Economic Indicators—a Progress Report* . 590
SIRKEN, MONROE, *Measurement Problems in Collecting Supplementary Data about Deceased Persons* 591
SNYDER, ELEANOR M., AND MYERS, JOHN G., *Methods of Inquiry—the Consumption Approach* 591
SOLOMON, MORRIS J., *A Decision Model for Computer Utilization* 591
SOMMERS, ALBERT T., *On the Logic of Diffusion Indexes* 592
STANBACK, T. M., JR., *Cyclical Behavior of Manufacturers' Inventories 1945–1955* . 592
ROBINS, MORTON AND STEINMAN, ALFRED M., *Ecology of Coronary Heart Disease: Statistical Implications* 592
SUNDELSON, J. WILNER, *How to Communicate Statistical Findings in the Field of International Markets to Management* 593
TERLECKYJ, NESTOR E., *Factors Underlying Productivity: Some Empirical Observations* 593
THOMAS, EARL A., *Some Estimates of Tolerance Limits from Censored Data* . . . 593
WALSH, JOHN E., *Approximately Distribution-Free $\bar{x}+ks$ Type Tolerance Intervals* . 593
WHITNEY, ALFRED G., *Characteristics of Life Insurance Sales and Buyers Which Influence Persistency in Premium Payment* 594
WILCOCK, RICHARD C., *Unemployment Effects of a Major Plant Shutdown in a Depressed Area* 594
*WILLIAMS, E. J., *Prediction of Clinical Effects of Rh Incompatibility* . . . . *Bio*
*WILLIAMS, JACK, *The Joint Action of Drugs in Tuberculosis* . . . . . *Bio*
WINE, LOWELL R., *Subject Matter Covered in Engineering Statistics at Virginia Polytechnic Institute* 594
WOODRUFF, RALPH S., *Methods of Estimating Retail and Consumer Inventories* . 594
*WOOLSEY, T. D., *The Development of Morbidity Survey Techniques for the National Health Survey Programs* . . . . . *Bio*

A Current Appraisal of Concentration Statistics. M. A. ADELMAN, *Massachusetts Institute of Technology.*

The new statistics compiled by the Census and in part released by the Senate Antitrust Subcommittee are a valuable addition to our stock of basic information. The statistics are sound in concept. The impression that they are in any sense inaccurate or invalid is incorrect. They do not measure the degree of monopoly, or even of fewness in an economic market; they form a body of evidence, greatly varying in precision, which does bear on the degree of fewness.

Concentration did not increase from 1947 to 1954. It does not necessarily follow that it will not increase in the future. The new statistics pose sharply the need for summary measures of concentration. To use measures of inequality would be a tempting but not a good expedient because inequality is not related to fewness and because the number of firms is unstable, and its meaning uncertain. Industry and product concentration statistics are complementary, not competitive. The chief problems in using them are of horizontal and vertical integration. Horizontal integration concerns substitution on the supply or demand side, and is not amenable to general statistical analysis. The "depth" aspect of vertical integration is subject to analysis and the chief tool available is some form of the ratio of materials purchases to sales. There appears to be wide intra-industry variation in this ratio. The identification of elements responsible for greater or lesser vertical integration, and of product flow, poses many problems, all expected to be illuminated by release of Census data now in process.

Rate of Change Approaches to Forecasting—Diffusion Indexes and First Differences. SIDNEY S. ALEXANDER, *Massachusetts Institute of Technology.*

A test of the predictive power of a current diffusion index based on fifteen to twenty-six of the components of the Federal Reserve Board Index of Industrial Production, as compared with the first differences of the FRB Index itself, leads to the following conclusions:

1. The *smoothing* operation is necessary but it may be performed currently with a loss in lead time that is small compared to the length of lead time involved. A currently smoothed diffusion index of the FRB Index components led business cycle peaks, in the interwar and post-war period, by 6 to 12 months.
2. The *differencing* operation is also fundamental but it may be performed currently on an aggregate, such as the FRB Index, about as effectively as on the components.
3. The *transformation* of the quantitative differences to a two-valued system of attributes: rising and falling, is not essential to the lead over turning points. The first difference of the FRB Index, which preserves quantitative relations, leads turning points just about as well as does the corresponding current diffusion index. A diffusion index based on a large number of series may be somewhat smoother than the first difference of the corresponding aggregate, however.

4. Nor is the *weighting* operation fundamental, since the equally weighted diffusion index and the unequally weighted first difference of the FRB Index behave similarly with respect to their leads over business cycle turning points.
5. While the use of *hindsight* is not fundamental to the lead over business cycle turns, hindsight *is* of fundamental importance in avoiding false forecasts, or more accurately, in distinguishing forecasts of minor fluctuations from those of fluctuations large enough to be called business cycles.

In short, a currently smoothed first difference of the FRB Index has about the same predictive value as a currently smoothed diffusion index based on the components of the FRB Index. Both lead all recognized business cycle peaks in the FRB Index (1919–1937, and 1947–1956) by 6 to 12 months, but both also reflect numerous minor intermediate fluctuations with a lead close to zero. The signals for a minor fluctuation do not seem to be contemporaneously distinguishable from those for a "real" business cycle turning point.

Some Aspects of Factorial Experimentation. R. L. ANDERSON, *North Carolina State College.*

Factorial experimentation was considered as to number of levels for each factor; nature of the levels (discrete or continuous) and corresponding experimental models; blocking arrangements (confounding); methods of estimating error variances; and methods of reducing the number of factor combinations used in an experiment (such as fractionalization). The Galois Field method of confounding and fractionalizing was discussed for 2 and 3-level experiments. For singly and fractionally replicated experiments, it was proposed that research was needed on the possibility of developing a sequential procedure to estimate the error variance.

Special consideration was given to the design and analysis of confounded factorial experiments, using orthogonal polynomial contrasts. The loss of information on first and second degree contrasts was derived for a number of confounded multi-factor experiments, using the usual confounded factorial arrangements (such as those based on Galois Fields). A new method of constructing such confounded designs based on the signs of elements in the orthogonal polynomial contrasts has been developed by Van der Reyden.

The recently developed response surface designs were discussed, with certain cautions regarding their general use. Finally split-plot experiments were mentioned, with an indication of the effect of interaction on the estimates and tests of hypotheses.

Measuring Labour Stability and Its Application to Employee Recruitment. J. E. ANDREW, *United Steel Co. Ltd. England.*

The increasing concern in industrial labour wastage is stressed, and an outline given of the principal objects of a labour turnover investigation. The inadequacy of the accepted measure of employee stability—the labour turnover rate—is illustrated, and alternative measures proposed involving the derivation of a survival function. Two methods of estimating this function are given yielding survival or stability curves. The approximation of these to a logarithmic normal form has facilitated the estimation of an expectation of service which is offered as a single measure of labour stability. Examples of the application of both methods to data collected in The United Steel Companies Ltd. are given. Research into factors underlying differences in stability between branches and between smaller groups of men is illustrated, and suggestions are made as to how this knowledge may be used in assisting labour managers towards labour budgeting and personnel selection. The age and living circumstances of the entrant in particular are shown to have a powerful effect upon the service that may be expected before resignation or retirement. The paper ends with details of proposed future studies into factors influencing labour stability.

The Implications of the Highway Program as an Important Factor in the Long-Run Economic Outlook. L. JAY ATKINSON, *Department of Commerce*

The Highway Act of 1956 provides a long-term stepped-up spending program for highways, with emphasis on interstate network and increased Federal share. Spending, including State matching funds on Federal Aid projects was about $1.7 billion in 1956, will nearly double within a few years and gradually rise to about $4 billion in the final stages. All of the spending will not be "net additional" spending. Toll road spending—now at a $1 billion annual rate—will drop and other projects in the planning stage may now be financed from Interstate funds. Strictly pay-as-you-build Federal financing will reduce the inflationary impact of the road program and tend to limit State borrowing. Estimated trust fund receipts available for Interstate spending may be inadequate for satisfactory progress on the Interstate program after the preliminary planning stage is passed. They are estimated to provide for Interstate spending of about $1.5 billion annually for the years 1960–65. Short term direct effects will be on materials, capital equipment, and skilled labor for highway construction. Longer term effects are stimulation of automobile industry, trucking, petroleum, and "roadside" group. Partial offset is in diversion of traffic and investment from other forms of transport, chiefly rail and transit. Interstate program will serve about one-fifth of traffic needs with fine roads, but other pressing traffic needs will be partly met by some increase in regular Federal aid and by rising State financing. For rural areas, traffic congestion is principally on interstate routes, and this will be largely relieved by the modern freeway system. More than half of the funds are allocated to urban areas. The urban roads will expedite a part of the city flow, but will attract more traffic. The net effect upon city traffic congestion will vary with size of city, transportation pattern, and other steps taken to relieve bottlenecks, especially the commuting jam.

Probability of Runs in Correlated Data. RALPH HOYT BACON, *General Precision Laboratory, Inc.*

Consider a sample of n specimens drawn from a "normal" population or universe, where the successive drawings are not statistically independent, and where n is a large number (say 50 or 100). A striking feature of such samples is the presence of runs (or sequences) of specimens all on the same side of the sample average. It is of interest to try to obtain an expression for the probability of runs of length r as a function of the correlation coefficients between all the pairs of specimens constituting the sample. As yet, the rigorous solution has not been obtained, but an approximation, which cannot be in error by more than 10% in the worst case, will be described.

The Demand and Supply of Statisticians from the Standpoint of Colleges. T. A. BANCROFT, *Iowa State College.*

The present and anticipated extreme shortage of able statisticians to fill academic positions in teaching, research, and consulting in statistics is, of course, part of the general growing dilemma in higher education. The over-all situation is spelled out in the recent report of the President's Committee on Education beyond the High School. Quoting from an editorial in the August 19, 1957 issue of the Washington Post, "The gist of this group's findings is that we are in an "explosive" period—population-wise and in the expansion of knowledge—and that the resources to meet the new demands for education are not at hand." The discipline of statistics, however, is in a particularly desperate situation due to the fact that it is academically a relatively new discipline as a separate field and at the same time finds itself in an "Explosive" period of development.

Modern procedures of statistical inference found early acceptance and use in the biological and agricultural sciences. In fact, many of the basic concepts and much of the basic theory and methodology were developed to provide research methods and techniques in the agricultural and biological sciences. During and subsequent to World War II industry, public health and medicine have discovered the great importance of statistics as a tool of research and management. Competition for the few students being trained in statistics in the few well established university departments of statistics by government agencies, industry, and the universities has reached an all time high and is mounting. Since the universities can be outbid even in competition for their own staffs, some devices must be found to retain able statisticians in academic positions to increase the supply of statisticians available for all. University positions should be made more attractive by making salaries more comparable with those in industry. Industry and government itself could help support the statistical programs at universities by establishing scholarships, fellowships, and supporting contractual research at the universities. The more intangible advantages of an academic career should be pointed out to able young statisticians, e.g. an intellectual environment, security, and freedom to engage in independent research.

Selected Stock Market Indicators and Their Current Implications. J. E. BANKS, *Brown Brothers, Harriman & Co.*

At the outset the limitations of any stock market indicators should be emphasized. Their value lies in their use as an aid in providing perspective for the longer-term investor. The relationship between stock prices and earnings is, of course, basic. This relationship may be examined by computing a regression between the Dow-Jones Industrial Average and the quarterly earnings on the stocks used to make up the Average. Current earnings may then be used to estimate a so-called "normal price." If the level of the Dow-Jones Industrial Average is below the "normal price" and earnings are rising, a favorable background for a rising market is suggested. The reverse is also true. An index which may aid in anticipating earnings trends is the ratio of bank debits to bank loans. This ratio is analogous to a ratio of sales to inventories but moves earlier. The debit-loan ratio also performs a second function in that it reflects developing financial strain or ease which may affect stock prices. A collateral indicator of the financial situation is the ratio of bank investments to bank loans. Yield relationships are also helpful at times. A ratio of the Dow-Jones Industrial Average yield to the yield on Moody's Aaa Corporate Bond Average has on occasion suggested longer term buying zones.

Regional Aspects of Cyclical Fluctuations. GEORGE H. BORTS, *Brown University.*

The analysis of regional business fluctuations has always provided an enigma to economists. Some have maintained that it is not worth while studying the cycle in regional slices of the economy; regions only mirror their composition of industrial or other economic characteristics. Others have dissented strongly in the belief that the regions have behavior patterns of their own which they impose upon the collection of national characteristics they happen to contain. I find myself thrust into this latter group by the findings of this study. The regions (in this case the 33 most important manufacturing states) do differ from each other in their cyclical behavior and differ in a way which the non-regionalists would deny was at all likely. On the other hand they differ in a way which the regionalists have not conceived. These differences are the subject matter of this report.

There are two interesting hypotheses in the field of regional fluctuations:

(a) One is that the cycle is transmitted to the region through a major, or export, industry. This means that the regional cycle reflects, or in some way, exaggerates the national industrial impulses transmitted to it. It is the reflection or exaggeration which constitutes the peculiarly regional element.

(b) The other hypothesis is that the severity of the regional cycle is in some way tied up with the growth patterns that the region experiences. One observer claims that cyclical severity is associated with growth. Another has claimed that cyclical severity is associated with economic decline.

It is of course not possible because of the scarcity of data to provide complete tests of the two general hypotheses. The tests which I have been able to make indicate the following:

(a) To the extent that the first hypothesis can be given content, it is rejected by the data in three out of the six cycles which I have studied. In addition it is rejected in a fashion which calls into question the entire hypothesis.

(b) Cyclical severity is not regularly related to growth, or to decline. It does appear related to retardation. That is, states with the greatest severity have simultaneously experienced a deterioration in their rate of growth relative to other states. Why this relation should hold is not entirely clear. Earlier studies of cycles in industries with differing growth rates have established that cyclical declines became more severe as the industry matures. The evidence of this study indicates that this phenomenon may extend in a more general form to state industries as well. That is, the over-all drop plus rise becomes more severe as the rate of growth retards in individual states.

Factors Affecting the Level of Fixed Investment in Manufacturing Industries for the Period, 1951–1955. MURRAY BROWN and HERMAN ROSEMAN, *University of Pennsylvania.*

We have attempted to explain the variation in fixed investment in manufacturing industries during the period, 1951–1955. For this purpose certain well-known theories of investment behavior have been tested—namely, the accelerator hypothesis, the profits theory, and a sales expectations theory (in lieu of actual sales). An introduction to the paper reviews the findings of previous empirical studies of investment (Tinbergen, Klein, Chenery, Clark,

Hickman, Modigliani, Koyck, Meyer and Kuh, and Norton), which served as a point of departure for the present study. Previous studies have presented conflicting reports, some indicating profits and some sales as the key explanatory variable. The present study compares these two variables (among others) as predictors and attempts to spell out some of the conditions under which one or the other is the better predictor. The data for this study is a subsample of 339 manufacturing firms taken from the SEC-OBE surveys of business investment. The SEC-OBE surveys provide data on firms' fixed investment and expected and actual sales; this has been supplemented by data from Moody's on profits, capital stock, and various financial variables. Our sample has been broken up by year and by industry; each industry-year group was analysed by linear multiple regression methods.

How An Economic Consultant Prepares Presentations for Clients. HENRY BUND, *Research Institute of America*

This talk centered on the fundamental differences between the presentation of statistical data for an individual company or client and that required in addressing a broad national audience or the members of a single industry.

In addressing its more than 30,000 members, the Research Institute is guided by these principles: the typical business executive today fully understands that his individual enterprise is closely tied to the future of the economy as a whole. But what he has learned to accept and look for are not facts but guidance, that is, interpretation of facts. What he wants is firm conclusions and recommendations, not a recital of data and reasoning which lead up to these conclusions. He expects to be alerted promptly to any significant shift in economic prospects, but does not want to be bothered by economic information when there has been no change.

In contrast to this appeal to a broad national audience, is the approach which must be followed with a more narrowly confined group such as the members of a trade association. Legal limitations preclude forecasting which might, conceivably, have a direct influence on an individual company's marketing or pricing decisions. As a result, much of an association's work has to be retrospective, designed to put the industry's performance in relation to the national economy and other industries. The most useful approach to this type of analysis lies in the presentation of industry-wide yardsticks by which each individual company can measure the effectiveness of its own operations with respect to each element of income and expense.

The function of both nationally disseminated economic guidance and the specific yardsticks available to the members of a trade association is to provide the background and general framework within which the analysts of an individual company can more effectively present their own conclusions to top-management.

Inventory Expectations Data. ORIN E. BURLEY, *University of Pennsylvania.*

It has been clear for some years that inventory investment constituted one of the most volatile types of business spending. Inventory status data frequently impart instability to the operations of the economy. The aim of this paper, however, is the consideration of the role played by inventory *plans* as related to understanding business fluctuations. In other words, when information on inventories is to be used as data foreshadowing economic change, the more pertinent data may well be *anticipation of inventory holdings* rather than historical holdings. One way to develop such a measure is to obtain information on anticipated inventory change directly from business firms. The determination of accuracy and timing should begin with a study of the past performance of the data treated as a time series. Conclusive evidence would demand that the series have a responsibility to continue in the future in essentially the same pattern that it exhibited in the past.

In its current experimental stage, the body of data available for consideration has no such full history. The experimental efforts of the Wharton School of Finance and Commerce study, the quarterly Fortune Magazine Survey of Inventory Plans, and the Dun and Bradstreet Survey of Business Men's Expectations do little more than suggest the potentiality of inventory plans as a forecasting medium. Nonetheless, on the basis of the implied efficiency of the data on inventory plans in forecasting the operations of the economy, the United States Department of Commerce has undertaken the collection of inventory plans data, on an adequate sample basis, which has considerable promise of adding a significant lead series to business forecasting knowledge.

Classification by Configuration and Level. E. I. BURDOCK, *New York State Department of Mental Hygiene.*

Classification of mental patients in terms of performance on a battery of psychological tests can be improved by reference both to configuration, i.e. pattern, of scores and to elevation, or level. If the test scores are drawn from a continuum with approximately normal distribution, the scores may be normalized. Differences in level may be tested by computing each subject's mean score and comparing the means of the groups within the population by the usual analysis of variance methods. It is then possible to test for differences in configuration by ranking the tests for each subject, group by group, and transforming the ranks to corresponding normal deviates, after which a test by subject analysis of variance may be carried out in which the mean square for the group X test interaction is compared to the error term. If this comparison yields a significant F-ratio, the groups can be said to differ in pattern. A three-way analysis may also be devised, as when the groups are divided among both sexes. The between-tests mean square may also be compared to the error term to see whether there is a significant over-all difference among the tests. However, this ranking method precludes an additional test of differences between groups, because the transformed scores necessarily sum to zero across subjects.

The above described method was used to analyse the test scores of 78 schizophrenic patients at Brooklyn State Hospital. The subjects were all between 18 and 50 years of age; 40 were males, 38, females. The subjects were divided into two groups, a chronic group which had spent more than 50% of the period since first admission in the hospital, and a mobile group with less than 50% of the time in the hospital. Nine psychological tests were used for the analysis: 1. A verbal analogies test, 2. A verbal test of essential differences, 3. Critical flicker frequency, 4. Reading color words, 5. Identifying colors, 6. Naming conflicting color of ink in which color names are printed, 7. Reading color words printed in conflicting color, 8. Tapping time, 9. Reaction time.

Results—Part I (Analysis of Means): F between groups =4.584; F between sexes =0.665 N.S.; F sex × group interaction =0.455. Part II (Pattern of 9 tests): F between tests =0.31 N.S.; F test × group interaction =0.37 N.S.; F test × sex interaction =4.90; F test × sex × group interaction =0.42 N.S.

Truncation to Meet Requirements on Means. F. EUGENE CLARK, *Rutgers University.*

The problem presented is that of truncating a normal population with known mean and standard deviation in order that the truncated population will satisfy requirements on sample means. The mean of a random sample from a lot submitted for acceptance must lie between specified values or the lot will be rejected. The procedure given here is to assign points of truncation in such a way that the truncated lot will meet the specifications with a prescribed risk of rejection. Cases of both single and double truncation are presented.

Anticipations Data in the Capital Goods Field. MORRIS COHEN, *National Industrial Conference Board.*

This paper represents a progress report on the new quarterly survey of capital appropriations conducted by the National Industrial Conference Board. The report sets the role of the appropriations survey in the postwar development of business statistics and the parallel trends in private business planning. It distinguishes between the appropriations approach and the surveys of capital spending intentions, emphasizing that the capital appropriation represents the company decision to spend for plant and equipment. It explains why the appropriations survey is limited to the 1,000 largest manufacturing companies, as measured by total assets. It outlines the battery of statistical indicators made available by the survey: appropriations newly approved, appropriation backlogs at the beginning and end of the quarter, appropriations that have been committed or spent, and cancellations. The published data for the ten quarters of 1955–1957 are analyzed briefly, the highlight being the apparent peak in new appropriations reached in the second quarter of 1956. Unpublished and preliminary data for 1953 and 1954 are discussed, in particular the fact that new appropriation approvals rose during the latter part of 1954 and early 1955 while capital spending declined. Finally, it is suggested that the appropriations series can serve as a short-term indicator, foreshadowing changes in manufacturers' spending for plant and equipment and providing perhaps the first clue to changes in such spending which may occur beyond the end of the calendar year.

Long Range Economic Outlook. GERHARD COLM, *National Planning Association, Washington, D. C.*

If reasonable full employment should be maintained, a Gross National Product of approximately $600 billion, measured in present prices, should be attainable in 1965. What are the existing forces operating in the economy which will bring about such economic expansion?

During recent decades, the most important forces making for expansion have been the following: 1. The technological innovations with their impact on business investments; 2. The change in the pattern of living—the suburban development with all its consequences for private and public outlays; 3. The high level of government expenditures for national defense; 4. The spread of installment buying; 5. The changes in the distribution of incomes—the rise in incomes of formerly low income groups, including the effect of old age pensions on consumer expenditures; 6. The rise in international trade activities. Some of these factors are now temporarily declining, others are rising or most likely to rise again in the near future. In addition, there is one group of demand factors which has been held back in the past but which is becoming of rapidly increasing importance. This refers to nondefense outlays of the government, such as for education, research, health, highways and skyways, conservation and development of water and other natural resources. Considering each of these various demand factors, the long range outlook for continued economic expansion appears most encouraging. There are, however, some factors which give reason for serious concern.

1. The unsolved problems of price and cost rises suggests that a restrictive monetary and credit policy may be continued. This policy which is designed to counteract price rises actually dampens expansion. Adequate and continued economic expansion requires a constructive reconciliation of the policies of price stabilization and of promoting needed expansion. 2. Most of the needed nondefense government services to which I have referred are in the area of primary state and local responsibilities. However, the tax sources which expand most readily with general economic growth are in the hands of the Federal Government. Thus, there arises the possibility that state and local tax resources may not be adequate to support the desired increase in government programs. This unsolved problem of Federal-state-local fiscal relationship presents a serious obstacle to those development programs which are of great importance for continued economic expansion.

3. Finally, there is a serious question about the future of the international economic outlook. The question of the dollar shortage which a few years ago appeared almost solved is coming up again. Continuing increase in exports will depend largely on a freer flow of capital from the capital-abundant to the capital-starved nations of the world. This flow is now impeded largely by political uncertainties. The conclusion is that the long-term economic outlook is most encouraging, provided that through public and private efforts we succeed in removing these serious roadblocks.

Some Experience with Electronic Computers in Processing Economic Census Data. MAXWELL R. CONKLIN and OWEN C. GRETTON, *Bureau of the Census.*

The paper discusses interesting and promising new developments in the use of electronic computers on economic census data by analyzing their application in the 1954 Manufactures Census. Both tabulation and mechanical editing phases of the processing are covered. In the tabulation end of the program, stress is laid on plant specialization statistics which were published for the first time in any census. The unique ability of the computer to manipulate data and compute ratios is described in detail.

The aids provided by the computer in detecting erroneous data are discussed. This section of the paper also deals with the 1954 Census program of mechanically imputing data for small establishments that fail to report certain items or reported them in an inconsistent manner. The electronic computer enabled the Bureau to test reported product and materials data against acceptable unit value ranges mechanically for the first time with significant savings in both time and money. In this part of the talk relating to products and materials there is also an explanation of the technique of spreading the tabulated establishment data in unit value frequency distributions for review purposes. Finally, some of the limitations of the use of the computer in 1954 are outlined in a section

which also describes the Bureau's plans for overcoming those deficiencies in the 1958 Censuses with the new scientific computer. The paper points out the vexing problem of introducing corrections into the computer tapes and related areas calling for improvement before the maximum advantages of the new equipment can be realized.

Inconsistency of Preferences as a Measure of Psychological Distance. CLYDE H. COOMBS, *University of Michigan.*

The measurement of inconsistency has long been a basis for measurement of psychological distance, as, for example, with the jnd concept in psychophysics. This experiment is an intensive study of the relation of measures of inconsistency to psychological distance in the context of *preferential choice* as distinct from the context of psychophysical discrimination. A theory was tested which predicts that measures of inconsistency of preferential choice are not monotonically related to psychological distance unless a new condition, called laterality, is held constant. A series of twelve shades of grey were presented in sets of 4 at a time to each of 4 subjects who indicated their preferences among the stimuli as best representing their concept of "grey." Each individual's concept of grey turned out to be an intermediate grey with lighter greys on one side and darker greys on the other. The hypothesis was tested that inconsistency between pairs of greys whose members were both on the same side of him would be of a different order of magnitude than inconsistency between pairs whose members came from opposite sides of him. Each subject constituted a separate experiment and the hypothesis was sustained for each of them. The significance of this for converting inconsistency measures into psychological distance measures was discussed.

Modifications of the Correlation Technique for Use in a Process Capability Analysis. DAVID COWAN, *Hughes Aircraft Co.*

Industrial processes are subject to myriad sources of variation, each of which may or may not influence the end product dimension. The first step in establishing control of a process is to identify and measure the sources which do influence the end product dimension. Correlation techniques, which potentially can both identify and measure have rarely been used in process control studies. The intricate and tedious computations connected with the usual multiple correlation technique are partly responsible for this situation, but there is a more basic cause. The technique does not yield practical measures in terms of a process, and the intricacies of the mathematical model do not lead to an insight into the physical relations among variables. The paper develops the statistic

$$P = 1 - \sqrt{1-\rho^2}$$

where P is the per cent of variation in the end product dimension that is associated with a variation in processing, ρ is the population correlation coefficient. An alternative multiple correlation model is proposed which can take advantage of existing knowledge of relationships among sources of variation.

Methods of Inquiry—The Income Approach. DANIEL CREAMER, *National Industrial Conference Board.*

This paper outlines the problems in measurement encountered in determining the extent and socio-economic characteristics of low-income families in New York State in 1949. The data for analysis were the income distributions of a one-per cent sample of families and unrelated individuals resident in New York State in April, 1950, and canvassed by the Bureau of the Census in the 1950 *Census of Population.* Attention is given to the development of an income norm that would mark off low incomes from higher incomes. The basis for setting up the norm is the standard urban workers budget necessary for the maintenance of health and decency. Its validity for our purpose is indicated by comparison with other budgets for other purposes. Adjustments to the reported income distribution were deemed necessary, particularly to allow for under-reporting of income and for improper classification of students living away from home. To come somewhat closer to a distribution of permanent income, an adjustment is made for the effect of cyclical unemployment on the 1949 income distribution. Applying our low-income yardstick to the adjusted 1949 income distribution, we estimate that about 30 per cent of the combined total of families and unrelated individuals had low incomes in 1949 and about 56 per cent of these had incomes below the public assistance budget.

A New Type of Estimate of the Income Elasticity of the Demand for Food. JEAN CROCKETT, *University of Pennsylvania.*

The well-known shortcomings of income elasticity estimates derived either from time series relationships or from Engel curves have led to a number of recent attempts to combine the two approaches in the hope of arriving at improved estimates. The present paper bases an estimate of the income elasticity for food on the changes over time in both food expenditures and income for a constant sample of about 1,500 families. Basic data were obtained from the records of the Market Research Corporation of America for the years 1951 and 1953.

If the family demand for food in the year t is a function of food expenditure, disposable income, family size, and other family characteristics affecting consumption, then an attempt to estimate the parameter relating to disposable income without the specific inclusion of family characteristics is likely to lead to biased results, since a fairly strong association exists between income level and family characteristics. However, if we consider a relatively short period of time, many of the family characteristics drop out since no change or negligible change occurs for most families in such factors as race, age, occupation, education, location, tenure of dwelling unit, and debt-asset position. Even when substantial changes in these variables do occur, it is not clear that they are highly correlated with income changes.

It should be noted that the functional form used implicitly assumes that there are no interaction effects either among the independent variables or between price and any of these variables (since the coefficients are constant over time rather than being functions of price). Prices are not necessarily the same for all families (e.g. they may be lower in small cities than in large) but are assumed to change in the same way for all so that changes in relative prices (as well as other factors acting uniformly on all families) are reflected in a constant term.

A shortcoming of the present approach lies in the fact that adjustments of consumption to large (and perhaps temporary) changes in income probably are not immediate and complete and may in fact be slower in some periods than others if the demand for other commodities such as housing and consumer durables is particularly strong. For this reason it was found desirable to replace current year income with a 3-year average income and to omit certain cases of unusually high income change. While this raises the estimates of income elasticity obtained from the 1951–53 data, these estimates are still somewhat lower than those usually obtained in cross-sectional studies. This suggests either that sufficient allowance has not been made for slowness in adjusting food consumption to income change or that the ordinary cross-sectional estimates are somewhat too high.

Accounting Records: A Limiting Factor in Economic Data Collection. EDWARD T. CROWDER, *Office of Statistical Standards.*

Economic statistics come to a large extent from business records. Survey planners often give insufficient attention to whether the facts they want are available in these records. Availability is not a simple matter, and unless researchers and respondents understand one another as to what a survey entails, there may be unfortunate consequences. The problem of availability may be illustrated in various economic fields. The researchers can turn to no comprehensive reference literature on what business records contain. He may investigate experience in similar surveys, consult business men singly or in groups, and make mail and field surveys of sources and availability of data—procedures which should be more systematically used. Better communication and cooperation are needed between economists and business experts on accounting and record-keeping. When data are, in effect, not available, abandonment of the plan in whole or in part may be indicated. Fragmentary data from a universe are dangerous to use, and research is needed on statistically valid procedures when records are inadequate for the whole universe. Businesses may be persuaded to keep special records for surveys, but the long-run solution is voluntary adaptation of the accounting procedures to the needs of economists. This can and will be done if business men and the community place a sufficiently high value on economic statistics.

Demand for Farm Products at Retail and the Farm Level. REX F. DALY, *Department of Agriculture.*

This paper isolates the influence of price and quantity variation on expenditures for food, computes price and income elasticities of demand, and relates the supply response to price, quantity, and expenditure elasticities. Expenditures for food have varied about in proportion to changes in consumer income. However, flexibility of real expenditure relative to real income is around 0.45, about the same as for cross-section surveys. The bill for processing and marketing food, around 60 per cent of expenditures, is somewhat more responsive to income changes than is total consumer expenditures. Value of products at the farm level showed a flexibility relative to income of 0.9 due mostly to variation in prices. However, income elasticity of quantity of food consumed at a farm equivalent level was only about 0.15—very unresponsive. A simple single equation demand analysis indicated price elasticity of demand for food at the farm level was about -0.12 and income elasticity was 0.18. A knowledge of the supply response is necessary to indicate how price, quantity and expenditure for food change with a change in income. With a very low price elasticity of demand, the greater the supply response to relative prices the lower the income flexibility of expenditure relative to income.

Factorial Experiment. CUTHBERT DANIEL, *New York.*

Fractional factorial 2^{p-q} experiments, with p factors, in 2^{-q} replication, are generally too large to run in one set. If all main effects and all two factor interactions ($2f$) are to be estimated, the numbers of runs required for p of 6, 8, 11, 15, are 32, 64, 128 and 256 respectively. The required fractional replicates are then 2^{6-1}, 2^{8-2}, 2^{11-4} and 2^{15-7}. *Sub*-fractions of each of these plans are given, requiring 16 runs, that provide estimates of large main effects, each confounded however with one or more $2f$s. Symmetric confounding patterns are given for these 2^{p-q-r} subfractions and for the corresponding experimental conditions for each run. The steps of augmentation to two-factor-interaction-clear plans are indicated.

Plotting of the e.c.d. of the collection of N-1 contrasts (main effects and interactions) produced by these experiments on a "half-normal" grid is proposed. Such plots facilitate detection of several kinds of defects, especially mavericks, inadvertent plot-splitting, and antilognormal error distribution, as well as aiding in the decision on what contrasts to pool for error-estimation. Also given are the operating characteristics of a "modulus ratio test" as found by A. Birnbaum and confirmed by some sampling experiments.

A Three-Sample Kolmogoroff-Smirnoff Test. HERBERT T. DAVID, *Iowa State College.*

A three-sample Kolmogoroff-Smirnoff statistic $D_{3,n} = \text{Max}\left[\sup\left\{F_{Y,n}(t) - F_{X,n}(t)\right\}, \sup\left\{F_{Z,n}(t) - F_{Y,n}(t)\right\}\right.$ $\left.\sup\left\{F_{X,n}(t) - F_{Z,n}(t)\right\}\right]$ is considered. The test based on $D_{3,n}$ is especially powerful against alternatives of type $X<Y<Z$ or $Y<Z<X$ or $Z<X<Y$. The distribution and asymptotic distribution of $D_{3,n}$ are, respectively, $\Pr\{D_{3,n} \geq l/n\} =$

$$3\sum_{i=1}^{[n/l]} \sum_{j \in J(i)} (\pm)(nl)^3/(n-il)!(n+l)!(n+(i-j)l)!,$$

and

$$\lim_{n\to\infty} \Pr\{n^{1/2}D_{3,n} \geq \lambda\} = 3\sum_{i=1}^{\infty} \sum_{j \in J(i)} (\pm)e^{-2\lambda^2[i^2+j^2-ij]},$$

where the set $J(i)$ consists of the integers $(2-i, 3-i, 5-i, 6-i, 8-i, 9-i, 11-i, 12-i, \ldots, 2-i)$, and where the $(\pm)$ sign indicates that, for fixed i, successive terms in the finite series indexed by j have alternating signs, beginning with $+$ for $j=2-i$, $-$ for $j=3-i$, $+$ for $j=5-i$, etc. The distribution is obtained by solving a random

walk on a triangular lattice in the plane, with triangular absorbing boundary; the solution involves a reflection principle in the plane.

Experiences with Response Surface Experimentation. ROBERT M. DE BAUN, *American Cyanamid Company.*

About one-half the work of the experimental statistics staff of a large chemical research laboratory is provided by fitting data to empirical response surfaces of the Taylor series type. It is urged that such work is not applied effectively when conducting "post-mortems" on previously gathered data or when the mathematical model of the physical system is even approximately understood. It serves best when the mathematical model of the system is only poorly understood, when fairly precise experimentation can be conducted and when an approximation to the true functional relation may be useful, either for empirical prediction or as an aid in arriving at an understanding of the system. Two examples are given, both of which show features characteristic of this type of experimentation. These are: multiple response variables, "ridginess," i.e., the absence of well-defined point maxima or minima, and the presence of additional non-continuous independent variables, such as block effects and discrete (qualitative) variables.

Tomorrow's Investments. HAROLD B. DORSEY. *Argus Research Corporation.*

The investment manager today cannot formulate his decisions on tomorrow's investment policies without coping with the question of inflation. Investment analysts have been seeing many practical violations of the theories that are supposed to support the idea of inevitable inflation, and, therefore, those theories are beginning to lose adherents. Noting the weakness in the prices of copper, lead and zinc, and in the earnings and stock prices of the producers of those metals, the theories of the wage-price spiral do not seem to be working out in practice. Higher wages are not being offset by higher prices, and the stocks of these natural resource companies are not functioning as a hedge against inflation. Nor are the theories of inevitable inflation working out for the railroads, or the airlines, or for the producers of building materials, textiles, home appliances, certain lines of paper, and numerous other segments of the economy.

It is quite clear to the investment analyst that some of the foregoing industries are now suffering from an old-fashioned over-expansion. We may not be much further along in the inflation cycle than many people suspect. The current debate about the inevitability of inflation fails to give enough consideration to the counter-inflationary effect that rising prices have upon consumption. The work that is being done by my staff suggests very strongly that administered wages and prices have not been able to overcome the simple and basic economic fact that rising prices tend to curb demand and encourage supply. The current inflationary condition is not a new, mysterious, economic malady for which there is no cure, and, therefore, investment policies should not be geared to the pattern of persistent inflation, with its extreme emphasis on investment in equities. The pattern of the investment program which seems most logical would allow for the possibility that we are going to continue to have business cycles.

On the basis of the fundamental economic forces, the investment manager should think in terms of carrying a balanced investment portfolio, i.e., balanced between good quality, fixed income securities and equities, and, in the latter group, his organization would be inclined to look with more favor on the defensive issues, until there is more assurance that the inflationary psychology of the last five years is not going to flop over to the deflationary side.

Communicating Statistical Findings to Management. STAHRL EDMUNDS, *Ford Motor Co.*

Economic analysis in the Ford Division provides part of the factual basis for planning on several specific operating problems. These are: production, product, financial, and facilities planning. The most continuous and exhaustive part of the statistical work is related to production planning, that is, planning production schedules throughout all the assembly plants and supplier divisions which will keep the output of cars in line with the retail market. The automobile industry operates with a very low inventory position compared with most other businesses. The amount of stocks of cars between the end of the assembly line and the consumer's garage is generally not more than 30 to 40 days' supply. This compares with an inventory of upwards of 60 days' supply in most other lines of business. To maintain a continuous flow of automobiles with moderate inventories requires that the production lines operate at approximately the same rate as consumer buying of new cars.

The cycle of product planning in the automobile industry is relatively long in that the product must be styled, engineered, and tooled as much as three years before actual production. The planning of a product which will appear on the market three years hence requires a close look at the kind, the number, the price ranges, and styling characteristics of the cars which are to be sold in the future. The relation of statistical and economic analysis to financial planning has two aspects. The first is to establish a reasonable and conservative base upon which tooling facilities, and product costs can be projected into the future. This requires the establishment of a standard volume of cars which can be sold and over which the heavy overhead costs can be spread. The second aspect of financial planning to which economic analysis contributes is in the area of pricing. Here the problem is to determine the volume of cars which can be sold at various price intervals. This helps determine the amount of cost which can be incurred and built into the product in relation to what the consumer wants and is willing to pay. The longer term projections of market conditions help to establish, also, the production facilities which will be required in the future. Hence, long range market projections help establish the expansion and investment policies which are appropriate to keep the company in a competitive position in relation to the total size of the market.

Bank Credit and Inventory Cycles. DORIS M. EISEMANN, *Rand Corporation.*

Changes in the availability or cost of bank credit may influence variations in business inventories. The proportion of inventories financed by bank loans increased during a cyclical inventory expansion and decreased during the contraction. At the peak of inventory accumulation, in the first quarter of 1951, almost one-third of manufacturers' inventories was financed through short-term bank borrowings. The correlation between the seasonally adjusted quarter to quarter percentage changes in loans and inventories for manufacturing between 1947 and 1955

was .91. During the three periods of credit restraint since 1951, the loan to asset ratio of small firms generally declined while that of large firms increased. As large firms also had a low correlation between changes in loans and inventories, we concluded that the level of their inventories would not be influenced by changes in the credit policy of banks. Inventories of small and medium sized firms were found to be sensitive to such changes.

Industries with large inventories in relation to their assets, and with a high ratio of loans to inventories, had the closest link between changes in loans and changes in inventories. Presumably, a change in monetary policy would influence the level of inventories of these industries more strongly than the inventories of other industries. For the period as a whole, the average magnitude of seasonal fluctuations was considerably larger than that of cyclical changes, for both loans and inventories. This points up one of the difficulties of using monetary policy to mitigate inventory cycles.

Karl Pearson: An Appreciation of One Aspect of His Work. CHURCHILL EISENHART, *National Bureau of Standards.*

An attempt is made to contrast briefly and as sharply as possible the differences in aim of *Biometry* on the one hand, and of *Adjustment of Observations* on the other. In a nutshell, the distinction is this. In *Adjustment of Observations*, variation is not a fundamental characteristic of the natural phenomena under study but creeps in through the imperfections of our instruments and our senses, and is thus something unwanted that has to be lived with. The aim of *Adjustment of Observations* is to minimize its influence. On the other hand, in *Biometry*, variation is recognized as an inherent characteristic of natural (at least biological and social) phenomena, and is neither to be ignored nor suppressed, and the aim of *Biometry* is to give proper expression to it and to measure it. Whereas the arithmetic of the Method of Least Squares, as developed by Legendre and Gauss, and the arithmetic of Correlation and Regression, as developed by Galton and Pearson, have much in common, the objectives of the two approaches are quite different; and Pearson's work is not merely an elaboration on the word of Gauss, but a tremendous step in an entirely new direction.

Presentation of Statistics to the Management of a Large Chemical Company. IRA T. ELLIS, *E. I. du Pont de Nemours and Company.*

The functions of the Economist's Office in du Pont are to study and interpret major business trends and developments. We report to the Executive Committee, the top operating committee of the Company. Our staff includes four men who are specialists in selected fields of economic research. One is an expert on commodity price trends; another is a specialist in the fields of economic theory and foreign trade. A third is a specialist in growth rates of particular industries. He is particularly familiar with the construction industry and population trends. Our fourth man has a background in chemical market research.

Our principal method of communicating economic information to management is through a meeting one hour in length on the third Tuesday of each month with the members of the executive and finance committees. Topics to be discussed are presented in the form of a report distributed the previous Friday. The entire hour is used for discussion. There is usually no formal presentation other than the report itself. The general theme about which our discussion subjects cluster is those factors that affect or are associated with economic expansion—the broad area of responsibility of these men—rather than the day-to-day activities of the Company. In addition to the above specific duties, we also prepare estimates of the probable business development one, or two, or ten years ahead, for use by the manufacturing departments in preparing sales forecasts covering various periods. Our estimates are usually restricted to such measures as the Federal Reserve Board Index of Industrial Production, indices of wholesale prices, and the consumer price index, wage rates in manufacturing industries, etc. These are usually transmitted in tabular form covering reported data for the past five years, together with estimates for the required future years.

While we prepare a variety of measures covering the entire economy, we also prepare a sales price index representing all our sales and one for each manufacturing department. We also prepare indices of "quantity" sales. Our commodity specialist maintains an index of market prices covering all the principal raw materials purchased by the Company. We also initiate special studies at irregular intervals, which are usually issued as formal reports.

Methods of Estimating Inventories of Passenger Cars. WILLIAM C. FLAHERTY, *Chrysler Corporation.*

The discussion of dealers' inventories includes the methods of preparing estimates by Chrysler Corporation, by the automobile industry, and by groups outside the industry. In addition, consideration is given to the usefulness, reliability, and comprehensiveness of the estimates. The treatment of consumers inventories is based on the number and age distribution of cars in use. Attention is paid to the usefulness of estimates for both short run and long run policies. Considerable care is paid to techniques of estimating future inventories both quantitatively and qualitatively. Mention is made of specific problems such as the impact of shifting scrappage rates for different year models.

Avoiding Nonsense Correlations in Time Series. RICHARD FOOTE, *Connell Rice and Commission Company.*

Nonsense correlations in economic analysis are those which reflect an historical accident rather than a basic economic cause. They may be dangerous for use in forecasting because the accidental relations that engender them can cease to exist at any time. More reliable forecasts usually are obtained from analyses that are based on the structural equations that are believed to make up an economic system, or their reduced form counterparts, even though such equations may give a lower coefficient of multiple correlation than can be obtained for other analyses during the period covered. Before using a study to make forecasts, charts should be made that illustrate the partial relationships between the dependent variable and each of the independent variables and the unexplained residuals should be plotted against time to determine whether they exhibit a trend or some other departure from randomness. Ideas obtained from an examination of such charts sometimes suggest sound modifications that need to be made in the analyses.

Interrelationships between Investment and Sales Anticipations. MURRAY F. FOSS, *Department of Commerce.*

Part I deals with the realization of business investment anticipations in 1955—a year which is of special interest in the study of factors affecting investment since it was a turning point year and one which saw a larger-than-average discrepancy between business investment expectations and results. The paper presents the results, cross-classified with respect to several variables, of a special questionnaire sent to manufacturing firms requesting explanations for deviations from 1955 plant and equipment expenditure anticipations as reported in the Office of Business Economics-Securities and Exchange Commission survey for that year. Results of relationships between investment deviations and deviations in sales and profits, both actual and anticipated, are also discussed.

Part II deals with factors affecting the investment anticipation (as distinct from actual investment) through a cross-sectional analysis of manufacturing firms. The paper presents, for the years 1954–57, relationships between expected investment, on the one hand, and ex post and ex ante variables on the other hand, such as sales, profits and depreciation.

Supply and Demand for Statisticians in Federal Service. DAVID L. FUTRANSKY, *Bureau of the Census.*

Discussion is restricted to those positions officially designated as Statistician. It will exclude all other groups such as Social Science, Economic, Mathematics, which may be applying and developing statistical methods but whose duties are not considered to be primarily in the field of statistics as defined in the Federal Classification system. In the past 10 years, the number of statisticians in the Federal service has increased from 1700 to 2100, with most of the increase taking place between 1947 and 1951. They are divided roughly as follows: Analytical Statistician—62%; Survey Statistician—15%; Statistician—15%; Mathematical Statistician—8%. (Mathematicians have increased during the same 10-year period from approximately 500 to 2000, with a 50% increase from 1954 to 1957.) In the nation-wide survey (June 1957), agencies reported a total of approximately 300 vacancies for which provision had been made in their budgets. Of these, at least 100 can be considered in the entrance grades of GS-5 and GS-7. Approximately 40 to 50 of the vacancies are for Mathematical Statisticians. The entrance grades of GS-5 and GS-7 ($3670 and $4525 per year) are covered by the Federal Service Entrance Examination (the old Junior Professional Assistant examination extended to include grade GS-7). This examination has not been able to provide enough available Statistician eligibles to meet agencies' needs.

The grades GS-9 ($5440 per year) and above are covered by unassembled examinations (rating based on review of education and experience). The registers for the Washington D. C. area are practically depleted of Mathematical Statisticians—less than 10 eligibles for all grades. Those on the registers for Analytical and Survey are mostly persons already working in the Federal service—of the approximately 320 persons on the registers, 75% are in the Federal service. The examination, therefore, acts more as an internal promotion service than as a source for bringing new people into the Federal service. Within the Analytical and Survey groups there is only a smattering of persons with subject-matter backgrounds in such fields as Demography, Health and Medicine, Physical and Biological Sciences.

The other big problem appears to be one of quality. In the past year some 215 certificates were issued in grades GS-9 to 15, of which about 25% resulted in an appointment. Apparently in many cases the certificates were returned unused because of the calibre of persons available for appointment.

How Results of Statistical and Economic Research are Presented to the Board of Governors of the Federal Reserve System. FRANK R. GARFIELD, *Board of Governors.*

One of the means of communication now used at the Board but not on display in Board publications is a visual-auditory presentation—"the show." This technique has been used since 1950 with Federal Reserve and a few other groups, domestic and international. "Each show," as noted in one of a dozen descriptive paragraphs, "has its own pattern, designed to fit the circumstances of the time and the particular audience, with due consideration also of what information, both new and true, may happen to have become available recently, from the ticker tape or a report on some research project. Whatever the pattern, a try is made to convey a sense of progression from introduction right through to conclusion—without being too dogmatic or too persistent about a particular theme. The natural inclination of responsible writers to avoid too firm commitments helps to maintain a scholarly tone of uncertainty. At the same time recognition of the necessity for decision making and of staff responsibility for sifting out technical perplexities does tend to bring definiteness into the picture."

The Low Income Problem: Basic Concepts. MEREDITH B. GIVENS, *Interdepartmental Committee on Low Incomes, New York State.*

The extent and causes of low incomes and depressed levels of living, which persist even in the midst of prosperity, are subjects of special current interest for research, social inquiry, and public policy. However, even in this statistical age we do not know with exactitude the absolute or relative size of the low income population. Perhaps of even greater importance we do not have a single, commonly accepted definition of low income. Administration and operations of the numerous public welfare and income security programs are producing a wealth of statistical detail about the recipients of each program. However, these materials are only of partial aid in developing a comprehensive picture of the numbers and characteristics of families and individuals at the lower end of the income distribution.

Clearly, a more complete description of the economic position and living conditions of population groups can be built from analysis which scrutinizes individual families. A primary source of such details is the decennial census. Unfortunately the decennial surveys do not provide any basis for other than national estimates. New York, like other states, is therefore handicapped in evaluation of current programs and in ascertaining the extent and type of unmet needs. An Interdepartmental Committee in New York State is taking action to fill in the statistical gaps in our knowledge of the size and characteristics of low income population. Its studies include (a) estimates of income distribution in relation to economic status, (b) a statewide survey of selected factors related to level of living,

(c) a survey of characteristics of public assistance recipients and their households, (d) an income-expenditure survey among households of unemployment insurance recipients, (e) analysis of standards of need in use by public agencies, and (f) study of problems of low income farming. Findings are to be used in framing an action program.

Application of Analysis of Variance Techniques to Profile Analysis. SAMUEL W. GREENHOUSE and SEYMOUR GEISSER, *National Institute of Mental Health.*

This paper does not consider the difficult problem of selection to form profile groups but rather the analysis of two questions often asked of group profiles for groups already formed on the basis of some external criteria. These two questions are: 1. Are the group profiles the same with regard to level? and 2. Are they the same with regard to shape?

Assume k tests (having the same metric) given to each of n_p individuals in the pth group ($p=1, \ldots, g$) yielding the following vector of means in the pth group: $(\bar{x}_{1p}, \ldots \bar{x}_{kp})$. Under classical analysis of variance assumptions, the most relevant of which is the independence (or at most equal correlations) among the k tests, it is intuitively shown that the appropriate statistic to measure departure from parallelism of the group profiles is the group means x test interaction. The appropriate error for this statistic is the individual x test interaction resulting in an F test with $(k-1)(g-1)$ and $(k-1)(N-g)$ degrees of freedom respectively, $N=\Sigma n_p$.

However, in almost all problems arising in psychology, the assumption of independence is violated and one must assume a sampling of vectors from a k-variate normal distribution. Based on an extension by Geisser and Greenhouse of results due to Box, the above test can be made approximately in the tabular F-distribution with $(k-1)(g-1)\epsilon$ and $(k-1)(N-g)\epsilon$ degrees of freedom where ϵ is a function of the elements of the variance-covariance matrix. A conservative test is an F with $(g-1)$ and $(N-g)$ degrees of freedom. The test on levels presents no difficulty and can be made with an F-test with $(g-1)$ and $(N-g)$ degrees of freedom. Problems of unequal covariance matrices and other relevant aspects are also discussed.

Mass Data Input to Computers, a Progress Report on FOSDIC. M. L. GREENOUGH, *National Bureau of Standards.*

For a number of applications the effectiveness of electronic computers is limited by the effort required for transcription of source data to a form suitable for computer input. Automatic equipment for converting data from marked documents to magnetic tape has recently been employed at the Bureau of the Census on a number of projects. Developed at the National Bureau of Standards, the equipment is called FOSDIC (for Film Optical Sensing Device for Input to Computers). In this instrument, microfilms of marked questionnaires are examined by an electronic, flying-spot scanner, the output of which is recorded on magnetic tape. FOSDIC is characterized by a high degree of tolerance for practical operating conditions, including provision for use of ordinary pencils or pens for sensing "dominant" marks and for freedom of document size and layout. In the paper, highlights of the present FOSDIC and new ones in process will be given. In addition to these mark sensing instruments, the recently completed, punched-card film reader will be described briefly.

Distinctive Characteristics of Psychometric Statistics—Problems of Metric. HAROLD GULLIKSEN, *Educational Testing Service.*

The metrics used to report test scores and other psychological "measurements" are open to many serious questions. Characteristics of various metrics used in the physical sciences are discussed with particular reference to the "addition experiment," from which these metrics derive their most valuable properties. There are analogous "internal consistency experiments" which provide similar valuable properties for certain psychological measurements. An illustration is given from the psychophysical method of paired comparisons, based on the law of comparative judgment. Multidimensional scaling methods and ability measurements method are also discussed.

Statistics for Engineers at N. C. State College. R. J. HADER, *North Carolina State College.*

Interest in statistics is growing very rapidly among engineering students at North Carolina State College. Recruiters from industry have been helpful in promoting this interest. Another factor has been the existence of a number of industry-sponsored senior year scholarships for engineers who agree to take additional work in statistics. At present one undergraduate course and several beginning graduate courses are available to engineers. The first course covers random variation, probability, a number of basic distributions, confidence intervals, comparison of two averages, tests of significance, simple linear regression and quality control. Quality control is also taught in several other departments. Attempt is made to emphasize concepts rather than develop skill in application. No mathematics beyond college algebra is used. The three graduate courses most popular with engineers have been Applied Mathematical Statistics, Basic Statistical Analysis, and Sample Survey Methods. There is also considerable interest in the IBM 650 which is located in the Statistics department.

Some Experiments with Demand Relationships for Clothing. MORRIS HAMBURG, *University of Pennsylvania.*

This paper reports some of the exploratory work in a study of demand relationships for clothing. The coverage is restricted to clothing expenditures as a whole although the completed study will concern itself with sub-groups and individual commodities as well. While the work that has been done and is reported in this paper is preliminary, the completed study will include a detailed analysis of the effects of a large number of social, economic, and demographic variables on clothing expenditures particularly as determined from the 1950 B.L.S.-Wharton budget study.

The first type of demand relationship for clothing that is derived is a least squares time series multiple regression equation relating per capita deflated clothing expenditures, per capita deflated disposable income and the relative price of clothing. Various transformations of the variables are made including first differences, logarithms, and first differences in the logarithms. A simple complete structural system is fitted to the same set of time series data, utilizing a marginal propensity to consume determined from the B.L.S.-Wharton budget study. Other at-

tempts are also made at the integration of cross-section and time series data for clothing expenditures. In all of the above-mentioned experiments, the emphasis is primarily upon the income and price sensitivity parameters obtained. Finally, certain aspects of one explanation of the decline in clothing expenditures in the post war period which relies heavily upon time series observations are tested with the B.L.S.-Wharton data. The major point of interest in these tests is whether, for urban families in 1950, low levels of clothing expenditures tended to be associated with high levels of durable goods purchases.

The Influence of the Interviewer on the Accuracy of Survey Results. R. H. HANSON and E. S. MARKS, *National Analysts, Inc.*

During the 1950 Decennial Census, the Bureau of the Census conducted an experiment to study the effects of interviewer variation. This experiment involved over 900 enumerators covering 1800 enumeration districts (ED's) with a total enumerated population of 980,000 persons in 24 counties in Ohio and Michigan. Assignments of ED's to enumerators were made at random within 134 geographic areas to control the effect of variation in the actual respondent characteristics and the interaction between respondent and interviewer characteristics.

Interviewer variance was analysed for 100 E.D. characteristics, including the usual Census characteristics (age, race, income, etc.) plus some special measures related to quality of performance (number of improperly omitted entries, difference of obtained from expected sampling proportions). As measured by the F-test, about $\frac{1}{3}$ of the 100 characteristics showed interviewer contributions to variance significant at the .01 level (at least one-third of these significant at the .001 level).

All of the measures of improperly omitted entries show interviewer variance contributions significant at the .001 level. Factors entering into significant interviewer contributions to variance appear to be (1) interviewer "resistance" to a given question—i.e., a tendency to omit or alter the question and to assume the answer; (2) relatively high ambiguity, "subjectivity," or complexity in the concept or wording of an inquiry; (3) the degree to which additional questioning ("probing") tends to alter initial respondent replies.

Also investigated was the relationship of interviewer performance to interviewer characteristics as given by (1) scores on a number of tests; (2) by such personal characteristics as age, sex, occupation, education; and (3) by attitudes toward, and expectations of, respondent reactions. At least one measure of poor quality of enumeration (number of improperly omitted census entries) seems to show substantial correlation with age, with some test scores, and with expectations about respondent cooperation.

The Demand and Supply of Statisticians from a Survey of Industries and Colleges. BOYD HARSHBARGER, *Virginia Polytechnic Institute.*

This survey covered a number of the large industries in America as well as various government agencies and colleges. The survey may be summarized in the following quotation from one of the respondents:

"There is a great shortage of well-trained statisticians who are taught how to use the theory. We could produce five times as many M.S. and Ph.D. level statisticians and not swamp the market in ten years. The recognized need is being expanded much more rapidly than production. An engineer who also has statistics training can be assured of $1,000 higher starting salary. An equally qualified statistician can easily secure $100 per month more than the same level engineer."

Maximum Likelihood Estimation from Sampling Data. H. O. HARTLEY, *Iowa State College.*

When estimating statistical parameters one is often faced with "incomplete data" in the sense that some of the observations in the random sample drawn are not available for the computation of the estimators. For example, if a time mortality study is terminated at a fixed point in time t^0, then the exact length of life will be known only for those experimental units dying *before* t^0, while for deaths occurring *after* t^0, only the number of deaths will be known. The estimation of the mean length of life from such incomplete data is known as "estimation from censored data."

If the number of deaths occurring after t^0, is not known either, we have what is known as a "truncated sample."

These are the best known instances of incomplete data and are contained as special cases in the more general situation dealt with in this paper, in which censorship and/or truncation may occur for any subset of the variate range. The computation of maximum likelihood estimators from such incomplete data is simplified and unified by a method which represents a considerable improvement on those in current use. It is in fact shown that maximum likelihood estimation from incomplete data can be reduced to that from a complete sample. The method is akin to the "missing plot technique" in analysis of variance, and it is applicable in any situation in which a maximum likelihood procedure for the complete sample is available.

The 1955 Survey of Life Insurance Ownership. ALBERT I. HERMALIN, *Institute of Life Insurance.*

The purpose of this paper is to analyze the economic and demographic factors which are related to life insurance ownership and the attitudes and motivations which affect life insurance behavior. The data are drawn largely from the 1955 Survey of Life Insurance Ownership, a sample survey of the population conducted for the Institute of Life Insurance by the Survey Research Center of the University of Michigan, although other studies of life insurance ownership are utilized.

The effects of income, age, place of residence, and education on life insurance ownership are set forth. Two patterns of life insurance behavior in the family are postulated—the tendency to have small amounts of life insurance on all family members versus the tendency to concentrate relatively large amounts on the family breadwinner. The income and educational characteristics associated with each of these patterns are presented. The manner in which the income and demographic factors are limited in their ability to differentiate large owners from small owners of life insurance is also discussed.

The importance of certain attitudes and motivations on life insurance ownership is pointed out, through findings relating the reasons people give for owning life insurance and people's feelings about the adequacy of their coverage to life insurance ownership. Differences in ownership apparent in certain demographic groups are seen, at least partly as differences in attitudes of these demographic groups. Possibilities for further research are suggested.

On the Problems of Coverage in Samples of Dwelling Units. IRENE HESS, *University of Michigan.*

An investigation of noncoverage of dwelling units in Survey Research Center surveys is presented. Formerly, noncoverage in our surveys was in the neighborhood of 10 per cent. Thus it was roughly the same magnitude as the total of all nonresponses, and its reduction was most important. During this period of low coverage, the dwelling unit addresses for SRC samples were selected from prelisted blocks and supplemented by the "half-open interval" technique at the time of interviewing. An analysis of the reasons for noncoverage showed that it was high in low economic blocks, in medium-size cities, and in both the slow growing and the fast growing places. In less than one year after new procedures were introduced into about 60 per cent of the sample, our coverage increased to approximately 97 per cent of the number of occupied dwelling units reported by the Bureau of the Census. These procedures are: (1) samples of city directory addresses supplemented by an area sample; (2) samples of small compact segments. Particularly gratifying is the success of city directory samples in the medium-size cities where coverage had been lowest. Now noncoverage, as a source of errors of nonobservation in our surveys, has been greatly reduced in magnitude and importance.

An Experiment with Weighted Indexes of Cyclical Diffusion. BERT G. HICKMAN, *The Brookings Institution.*

Diffusion indexes measure the extent of common participation of the various sectors of the economy in expansions and contractions of aggregate activity, and are used to study the relationships between the degree of participation and other characteristics of business cycles, such as turning points and amplitudes. The usual index takes account only of the direction of change of the components, so that if in a group of 100 time series, 80 were increasing and 20 decreasing in a given month, the value of the index would be 80 per cent. An experiment was undertaken to determine if the behavior of a diffusion index of employment would be affected when it was weighted to allow for the size of the component industries, the amplitudes of the individual fluctuations, or the combined influence of size and amplitude. Size weighting made little difference, because the average size of industries in the expanding and contracting groups remained fairly constant despite substantial fluctuations over time in the number of industries contained in the two groups. Amplitude weighting exaggerated the monthly movements of the unweighted index, because the forces which made for an increase of diffusion also tended both to amplify the average increases in rising industries and to dampen the average declines in falling industries. Some suggestions are offered about the use of diffusion indexes in the search for the causes of business cycles.

The Design of an Experiment to Measure the Effectiveness and Costs of Programs of Services for Aging Persons. JULIUS A. JAHN, *Community Service Society of New York.*

The existence of relatively high rates of unemployment, low income, illness, disability, sub-standard housing, social isolation, personal dissatisfaction and similar "problems" among aged persons as compared to younger persons has become a matter of common knowledge and concern in many communities. A great variety and number of programs have been undertaken or advocated to "help older persons" affected by such problems, and large expenditures of time, money and other resources have been made and would be required to carry out such programs.

Since resources for such expenditures are limited, a "good principle" to adopt in making decisions concerning which programs should be developed and administered is to select from all possible programs that one or few which will be "most effective" in overcoming specified "problems" of aged persons within given limits of "cost." The determination of the "effectiveness" and "costs" of possible programs involve hypotheses that are subject to errors, so that in practice the programs selected may not actually be the "most effective" for given costs. However, by the use of acceptably designed experimental research methods, some of these errors can be eliminated and others can be reduced to specified limits of size and probability of occurrence.

Accordingly, the Community Service Society of New York City has been concerned not only with the development and administration of health and welfare services for older persons in New York City, but also with planning and carrying out experimental research to measure the "effectiveness" and "costs" of a number of alternative programs. The design for such an experiment is presented in this report.

An Appraisal of Productivity Projections. JAMES W. KNOWLES, *Joint Economic Committee, United States Congress.*

The tests suggested in a previous paper are applied to the appraisal of productivity projections, of which 25 are summarized in tabular form. First, were the projected quantities the best estimates which could be derived? This could not be answered with confidence, except in a minority of cases. Second, would other assumptions have been more efficient or appropriate for the particular purpose? Given the structures of the mode, in most cases changes in the productivity projections would not have improved the over-all economic projections, nor would mere changes in assumptions have improved substantially the productivity projections. Third, in most "full employment" projections, the structures of the models were not optimum for the purposes leading to their construction. Two reasons for deficiencies in these productivity projections were found. Inadequacies of both data and empirical research into the basic relationships between inputs, outputs, prices, and factor returns hamper the makers of such projections. Also, many mistakes in such projections can be traced to failure to start with a clear and precise statement of the objectives or purposes of the projections.

Farm Income Estimates—Methods and Meaning. NATHAN M. KOFFSKY, *Agricultural Marketing Service, USDA.*

The Department publishes regularly estimates of gross farm income, production expenses and net income of farm operators. Net income is shown for the U. S. and by States on a realized basis and on a net income basis including the value of inventory change. In addition, the Department publishes income of farm people including income from nonfarm sources and comparisons of incomes of farm and nonfarm persons. The estimates of gross farm income are based largely on information collected by the Crop Reporting Board supplemented by the Census of Agriculture every 5 years. However, the information on production expenses and on nonfarm income received by farm people has been relatively weak. These have been reinforced by a recent nationwide survey of farmers' expenditures and off-farm income. For the future, estimates of farm income by economic class of farm are being developed, particularly the break between commercial farms and noncommercial farms. There is also the continuing problem of how to assess differences in incomes between farm and nonfarm people. This involves the valuation of nonmoney income received from the farm, the return to farm investment, differences in the cost of living as between farm and city, relative skills and other factors.

The Contribution of Cross-Sectional Survey Data to Prediction of Aggregative Components of Consumption. HELEN HUMES LAMALE, *U. S. Bureau of Labor Statistics.*

Only a few years ago consumption functions for different types of goods and services were called the "wall flowers" of economic analyses, while search for *the* consumption function centered in time series analyses. The interrelatedness of time series and cross-sectional survey data and the need for understanding the component parts of aggregate consumption have been recognized. However, consumption functions derived from time series analyses and cross-section analyses differ. It is in the explanation of these differences that cross-sectional survey data make their greatest contribution to prediction of aggregate consumption.

The cross-section studies provide detailed information for various components of aggregate consumption in relation to the demographic and economic characteristics of the consumer. These data, when properly analyzed, will provide many of the links to time series analyses needed for predictive purposes. Proper analysis, however, involves consideration of differences in concepts, definitions, and classifications between the time series and cross-sectional data. Some of these differences are discussed. Thus, the cross-sectional data provide the raw ore from which prediction tools may be derived but only after careful refining and processing.

Median Age at First Marriage in the United States. EMANUEL LANDAU, *California Department of Public Health* and PATIENCE LAURIAT, *Bureau of the Census.*

The medians discussed in this paper were calculated from decennial census data on marital status by single years of age. The tables (to be distributed) present estimates for 1940 and 1950 of median age at first marriage for States by sex, residence groups, regions, and selected States by color and sex.

The most striking characteristic of the rapid decline during the 1940's in median age at first marriage is the significant lessening of differentials between the various groups in the population. Generally, the groups with the highest ages at marriage showed the greatest declines. The decline between 1940 and 1950 for white persons was about one and one-half years, whereas the decline for nonwhites was less than half a year. The decrease for urban areas was almost twice as great as that for rural areas, for both whites and nonwhites.

The Northeast, the region with the highest age at marriage for both males and females in 1940, shows the greatest decline during the 1940 decade, reflecting this region's large proportion of population living in urban territory. In each region, with the exception of the West, where in 1940 the median age at first marriage for nonwhites was higher than for whites, the decline among white persons was almost twice as large as that among nonwhites. In the West, the greater decline among nonwhites reflects the sharp change between 1940 and 1950 in the racial composition of that region's nonwhite population.

The tendency for groups with the highest ages at marriage in 1940 to decline the most between 1940 and 1950 is also evident in the data for States. Furthermore, the data indicate a striking similarity between States in the difference between the median age at marriage for males and females in both 1940 and 1950. The data for selected States also show a more rapid decline for whites than nonwhites, as noted above.

A brief discussion of the method used in computing these medians is included, together with a short statement regarding differences between the various median ages at first marriage which are found in the literature.

A Model for Failure Data. ANDRE G. LAURENT, *Wayne State University.*

Given a "stimulus" S with cumulative effect X and a "System" $\sum$, the response of the system can be described by the "X-to failure" probability distribution. The exponential type of failure which implies the constancy of the expected life of the system and is valid when there is no ageing process and death is accidental, has been thoroughly studied. A model is proposed which is characterized by an exponential law of decay for the expected life or "resistance to failure" of the system, namely $E_x = E_0 \exp(-aX)$. To estimate the parameters of the model, least squares, maximum likelihood, moments methods, and methods based on ordered statistics are studied. The cases of truncated and censored samples are considered.

Use of Sample Survey Data and Analysis of Variance as a Guide to Packaging and Pricing Policies. DICK A. LEABO, *University of Michigan.*

The purpose of this paper is to demonstrate the usefulness of the application of the techniques of analysis of variance to sample data as a guide to packaging and pricing policies. While the data employed are hypothetical, *actual* market information may easily be substituted and the procedure remains the same. Therefore, while the conclusions based on these hypothetical data *may* not be true of a given market situation, the technique is still valid.

Evaluating Influences on Satisfaction with Patient Care in Hospitals through Multiple Regression Techniques. EUGENE LEVINE and BENJAMIN BUCHBINDER, *Public Health Service.*

While the number of professional nurses in hospitals has grown to 265,000 and the number of non-professional or auxiliary personnel has reached nearly 400,000 there continues to be an acute and growing shortage of nursing services. However, evidence has been piling up to the effect that while a shortage of nursing *services* may exist, it is not entirely due to a shortage of nursing *personnel*. Some hospitals with high daily average hours of nursing care available per patient complain as loudly about being understaffed as hospitals with low hours. It seems as though factors other than the number of nursing personnel, such as poor utilization of nurses' time and high personnel turnover, generate feelings of a shortage of services and consequent dissatisfaction with nursing care. The ratio of *professional* to *auxiliary* personnel on the nursing unit and the size and type of control of the hospital are other possible influences on the amount of nursing services provided and the level of satisfaction.

In the spring of 1956, the Public Health Service, in cooperation with the American Hospital Association, launched a study in 60 general hospitals to explore the relationships between the number of hours of professional and auxiliary nursing care available per patient per day, the size and type of control of the hospital, and the feelings of shortages of nursing services, from the viewpoints of the patients, doctors and nursing personnel. Checklists were developed on which 20,000 patients and personnel recorded actual omissions of important components of nursing care. The items were weighted by their importance, and average total scores of satisfaction with nursing care were computed for each type of respondent in each hospital. A separate multiple regression analysis was carried out for each type of respondent. In the model, the dependent variable was this average total score of satisfaction expressed as a function of unknown parameters representing the amounts of professional and auxiliary nursing care available and the size and type of control of the hospital.

The findings of the study indicate that more can be done to alleviate the shortage of services and the concomitant dissatisfaction, than merely recruiting more personnel. For example, patients' satisfaction is related to the number of hours of professional care available, but not to the number of total (professional plus non-professional) hours available. This is a strong indication that not just care by *any* nursing personnel is important to the patient, but that the skills and knowledge of the *professional* nurse enable her to understand the patient's problems and therefore give care which transcends merely meeting the immediate physical needs.

The Role of Industrial Classification in the Employment Security Program. LOUIS LEVINE, *Bureau of Employment Security, U. S. Department of Labor.*

The Bureau of Employment Security and the affiliated State employment security agencies have need for a wide range of data classified by industry for use in day-to-day operations, administrative planning, and economic analyses. The Federal-State employment security system encompasses programs for job placement and employment counseling, stabilizing employment, minimizing unemployment, reducing labor turnover and providing income maintenance through partial compensation of wage loss resulting from involuntary unemployment.

Employment and earnings of workers covered by unemployment insurance laws, insured unemployment, estimates of total employment, labor turnover, employer estimates of future labor requirements, job openings listed by employing establishments, and workers placed in jobs by the public employment offices illustrate some of the types of data developed and utilized in the employment security system. Effective use of these data make it necessary that they be classified by the State agencies uniformly by industry. These data serve important purposes for others, both within and outside government, for program, marketing and general economic analysis. The revised Standard Industrial Classification system will make these data more significant and more useful to the system and to other users by facilitating comparison and coordination with data from other sources similarly classified.

Variations in Mortality from Heart Disease. EDWARD A. LEW, *Metropolitan Life Insurance Company.*

The paper presents the variations in the recorded mortality from arteriosclerotic heart disease, including coronary disease, in the United States with respect to age, sex, race, marital status, geographic region, socio-economic status, occupation, and a number of other factors. Corresponding data for Great Britain and other countries are also cited where available. It appears that many factors, seemingly unrelated, affect the death rate from arteriosclerotic heart disease. If this is true, one should guard against the tendency to over-simplify and over-generalize in accounting for the variations in the incidence of the disease. The hypothesis of multiple causation appears to be most consistent with the known facts. The paper also calls attention to some of the gross biases that exist in the recorded mortality statistics for degenerative heart disease. Some of the difficulties inherent in the interpretation of such statistics may be overcome if greater uniformity can be secured in the terminology used in reporting heart disease on death certificates and if analyses of mortality from heart disease are extended to give consideration to multiple causes of death and to histories of morbidity.

Table for the Determination of Two-Sided Tolerance Limits for the Normal Distribution. ALFRED LIEBERMAN, *Dept. of the Navy.*

A solution for the computation of tolerance limits for the normal distribution has been given by Wald and Wolfowitz. Based on these results Bowker has prepared a table of factors K such that the probability is γ that at least a proportion P of the distribution will be included between $\bar{x} \pm Ks$, where $\bar{x}$ and s are estimates of the mean and the standard deviation computed from a sample of N.

The table prepared here is an extension of Bowker's table. Bowker has determined K factors for means based on samples of size N and standard deviations based on $N-1$ degrees of freedom. The present table gives K factors for means based on samples of size N and standard deviations based on n degrees of freedom. A special case of the present table is the case when $n=N-1$, in which case the present table gives the same results as those given by Bowker.

The K factors are given for $\gamma = 0.75$, 0.90, 0.95, and 0.99; for $P = 0.75$, 0.90, 0.95, 0.99 and 0.999; for values of N from 1 to 10, 20, 30, 50, 100 and ∞; and for values of n from 1 to 30, 40, 50, 100 and ∞.

The primary additional use of the present table over those already existent is that it can be used in conjunction with experimental designs, where the number of degrees of freedom determining the standard deviation is usually not $N-1$. Another application of the present table is in the setting of tolerance intervals for regressions where the number of degrees of freedom is not $N-1$.

The Contribution of Panel Studies to Forecasting the Automobile Market. CHARLES LININGER, EVA MUELLER, and HANS WYSS, *University of Michigan.*

Periodic surveys, each conducted with a new representative sample of consumers, indicate a close relationship between changes in aggregate car buying plans and changes in total sales—with only occasional exceptions. Reinterviews of the same people are a powerful method whereby forecasts can be checked on an individual basis as well. At the same time they enable us to assess the predictive value of additional variables which may be thought to influence automobile demand. They can also show what intervening changes are responsible for non-fulfillment of plans or unplanned purchases. In addition, panel studies may provide purchase data for the same families over an extended period of time.

The principle source of data for this paper is a panel study by the Survey Research Center financed by the Ford Foundation. It consists of an original interview in June 1954 with 1150 urban families and 4 subsequent reinterviews, the last one in February 1957. In addition the paper draws on data from reinterviews with the Survey of Consumer Finances sample in 1949 and 1953.

One part of this paper presents data on the pattern of automobile buying over a 3½ year period within groups differing by age of the family car. Car purchase plans, their fulfillment, as well as unplanned purchases also are examined within these groups. A second part of this paper compares fulfillment of car buying plans in relation to certain characteristics of the plan—plans to buy new cars vs. plans to buy used cars, plans expressed with certainty vs. plans that are uncertain, short-run vs. long-run plans, planned cash purchases vs. planned credit purchases, and so forth. The third part investigates the contribution which attitudes other than buying plans may make to forecasting automobile purchases.

Research Methods for Analyzing Short-Term Demand for Consumer Durables—Housefurnishings and Equipment. VERNON G. LIPPITT, *General Electric Company*

This paper deals with the question: What research methods show most promise of yielding an understanding of consumer demand for housefurnishings and equipment—an understanding adequate for forecasting aggregate demand for these items reasonably accurately for twelve months into the future?

Research methods discussed include: extrapolation; correlation of aggregate time series to predict aggregate demand; use of aggregate time series to predict segments of total demand; surveys of intentions to buy; use of cross-section data to determine relation between buying behavior and household characteristics; combining evidence from cross-section data and time series, at the aggregate level or through repeated cross-sections or reinterview data. Problems of aggregating the effects of family characteristics are discussed. Suggested areas for further research effort include: more analysis of existing cross-section data, probably using analysis of variance approach; further efforts to combine evidence on effects of family characteristics and market conditions, especially through repeated surveys or reinterviews; investigation of the stability or trends of demand relations through time; experimentation with Orcutt's model involving a sample from the universe of consumer units and direct aggregation of buying behavior; securing more survey data on family expenditures by major categories, including longitudinal studies of given consumer units over time.

Some Statistical Problems in Urban Renewal. BENJAMIN LIPSTEIN, *Market Planning Corporation.*

The planner's needs include both bench mark data on the status of the area and information on trends. The sample survey of population and housing is capable of producing the basic bench mark data with a high degree of reliability and simultaneous information on trend movements for the population that resides in the area. There are, however, a number of unique and difficult problems to be considered in conducting surveys in urban renewal areas. (1) The standard definitions should be examined for their applicability to local conditions. (2) Consideration must be given to proper public relations, lest refusal rates make the study useless. Both the size of the sample and the questionnaire mechanism must be geared for adequate control of interviewer variability and application of definitions which are more complicated than those usually found in consumer and market studies.

Automatic Programming as a Means of Enabling Statistical Applications of Computers. MITCHELL O. LOCKS, *Sperry Rand Corp.*

The statistician encounters some difficulty in using automatic high speed digital computers. To program an application, it is necessary to consume a great deal of time in applying an intricate system of codes. Automatic programming systems can reduce the barrier between the statistician and his computer. The most advanced type of automatic programming is the compiler. Under a compiler system, pseudo-codes are used instead of machine codes. Each pseudo-code refers to a complete subroutine in a library of subroutines. By using a compiler's simple and mnemonic pseudo-codes, the application can be programmed in a small fraction of the time that it can be directly in machine codes. The empirical portion of this paper deals with a case study in using the A-2 compiler system of UNIVAC I in coding a General Purpose One-Way Analysis of Variance. After verification of formulas, analysis, and flow charting it took two hours to code this problem in 26 steps of pseudo-code. This resulted in 600 machine instructions. Compiling time (i.e., the amount of time taken by the machine to write its own program) was about one minute. Running time was negligible on some test problems.

Distinctive Characteristics of Psychometric Statistics—Problems Arising from Errors of Measurement. FREDERIC M. LORD, *Educational Testing Service.*

Since observed scores on mental tests contain sizable random errors of measurement, the basic problem of psychometric theory is to obtain valid inferences about true scores from the scores actually observed. Quite satisfactory assumptions about the frequency distributions of the errors of measurement are available, and the psychometrician can obtain experimental data from which the parameters of these distributions can be estimated rather accurately. Unsolved (or inadequately solved) but very basic problems include: (a) estimating the true score of each individual in the group tested from his observed score; (b) estimating the shape of the frequency distribution of true scores for the group tested; (c) testing the hypothesis that two different tests measure the same thing, i.e., that their true scores have a perfect curvilinear correlation; (d) estimating the shape of the functional relation between true scores on two tests, or of their regression curves. Each problem is outlined in mathematical form, together with the kind of data available and a possible set of appropriate assumptions.

The Evidence Relating Diet to Coronary Heart Disease. GEORGE V. MANN, *Harvard University.*

The origins of the hypothesis relating the human dietary [illegible] coronary heart disease (CHD) date back to the eighteenth century. The relationship became of theoretical importance in the 1850s when Virchow and Rokitansky debated the pathogenesis of the fatty, arterial lesions they observed. The dietary causation of atherosclerosis was emphasized by the production of atheroma in rabbits by dietary regimens. All subsequent experimental atherosclerosis has been accomplished by increasing the serum cholesterol level and this has generally been done by dietary means.

We can not conduct the critical human experiments that would be the direct approach. These are impossible for ethical reasons and also because of the chronicity of the process we are studying. It seems important to relate the usual diet to the level of serum lipids, to atherosclerosis—an anatomical process—and to the behavior of CHD —a clinical phenomenon.

There are four ways to evaluate dietary intakes (1) food disappearance data gathered by economists (2) family surveys (3) individual dietary records (4) personal interview estimates. The techniques and the limitations of each will be discussed.

There are no direct methods for evaluating any of these in an absolute manner. The indirect ways consist of (1) correlated evaluation of subjects' physical status (2) comparisons of energy intake measured by the dietary appraisal with estimated energy expenditure (3) small scale samples of food weighing data. Experience with these methods will be discussed.

The enticing prospect of a quick and simple dietary evaluation by application of a short questionnaire designed to measure the intake of a few "key" foods has not been successful because the evidence for identifying these key foods—if such exist—has not been obtained.

Some Probability Considerations in the Estimation of Computer Running Time. ELI S. MARKS and BENJAMIN J. TEPPING, *National Analysts, Inc.*

When the operations performed by a computer are carried out serially the estimation of computer running time is quite straightforward, running time being the sum of the time taken by the various individual operations, each of which may be estimated separately. However, one of the major features of modern high speed digital computers is that they are designed to read, write, and compute simultaneously or nearly so. Where the computer program involves handling relatively large volumes of input and output with a moderate amount of manipulation of the data internally, the problem of estimating total time required may become difficult.

One way of estimating time in such circumstances is to program the problem in detail and observe the actual running time. This may not be possible since (a) the equipment proposed may not exist at the time that the estimates are required; (b) it may be very costly and time consuming actually to obtain fully tested programs. The procedure described in this paper is first to construct a program which is not necessarily complete or accurate in all details, but is sufficiently so as to provide reasonable estimates of the time required for the basic operations that are involved in the program. A probability model is then set up, on the basis of assumptions regarding the distribution of the input by "type," the latter being defined by the basic operations that are required to handle the input record. On the basis of the probability model it is possible to estimate the means and variances of the delays that are encountered, that is, the intervals of time during which operations are interrupted because other operations must be completed first. This leads directly to estimates of the expectations and variances of the elapsed time for handling records drawn at random from a specified statistical population. An application is described in which the relative efficiencies of two computers are compared, for a given file maintenance problem. For one of the computers, the application involved "double-programming."

Greenhouse and Laboratory Experiments. CARL E. MARSHALL, *Oklahoma State University.*

The sources of variation, limitation of facilities, and the techniques of caring for experiments cause greenhouse and laboratory experiments to be as hazardous as any other type of experimentation. He who feels that the variation in this type of experimentation is so small it can be ignored, should experience an appropriately designed uniformity trial. A critical look at greenhouse and laboratory experiments reveals that although the experimenters are interested in testing hypotheses about central tendencies and measures of variation, they have considerable interest in the areas of response curves and estimation.

Present day experimenters are discarding the one factor approach in favor of designs which will accommodate the more natural conditions, where many factors vary simultaneously and frequently not independently. Factorial arrangements of treatments are proving quite valuable. The use of the Latin Square, the Youden Square, and other balanced and partially balanced incomplete block designs are increasing the efficiency of greenhouse and laboratory experiments. This type of experimentation usually implies a simulation of conditions or situations which are ex-

pected to be experienced on a much larger scale. The statistical inference must of necessity be made to the population of experimental units and any transfer of inference to any other population of interest must be based on judgement, logic, or intuition. An exploratory experiment should be interpreted as such. Frequently it will be advantageous to disregard the principle of fixed sample size and turn to the multi-step approach of experimentation. This sequential type of experimentation is especially useful in laboratory experiments.

The statisticians should be urged to publish articles concerning statistical techniques in the language of, and through the media of, the various subject matter fields. One method open to all is to work for joint publications in the fields which demand their attention.

The Predictive Value of the National Bureau's Leading Indicators. FRANK E. MORRIS, *Investment Bankers Association of America.*

The National Bureau's statistical indicators were not intended to constitute a complete basis for forecasting. In themselves, they provide the foundation for only a partial analysis of cyclical trends. The user must modify the results given by the indicators in the light of other information available to him. This fact complicates any attempt to make an objective appraisal of the predictive value of the indicators during the postwar years.

In general, the lead-lag relationships which prevailed during the prewar years have continued to prevail during the postwar years, despite important structural changes in the economy. A competent analyst, using the National Bureau indicators, should have been able to produce reasonably accurate forecasts of the recessions and revivals of 1948–49 and 1953–54. It is possible (although certainly not inevitable) that he may also have predicted fluctuations which never occurred (at least on any significant scale). The user of these indicators during the postwar years may have made mistakes in forecasting the timing of cyclical turning points and in judging the amplitude of swings in the economy. He should not, however, have made mistakes with regard to the direction of movement of the economy.

A number of difficult problems are encountered in applying the indicators to forecasting on a current basis. The most important of these are: (1) erratic movements in the basic data, which cannot be entirely eliminated by smoothing devices; (2) substantial differences in lead times from cycle to cycle; and (3) the fact that the leading series is responsive to leveling-off periods or minor cyclical movements as well as to major cyclical trends.

The difficulties caused by erratic movements in the basic data can be reduced by introducing a new dimension in obtaining a consensus of the indicators. The consensus measure most commonly used is the "percentage expanding" index, which merely records the direction of movement of the indicators. Another measure in use is the "duration of run" index, which weighs the direction of movement of the indicators by the number of consecutive months that the movement has occurred. This produces a smoother index, but one which occasionally lags behind the "percentage expanding" at cyclical turning points. Neither of these commonly-used indexes reflects the rate of change in the indicators.

Experiments with consensus measures which give weight to the rate of change in the indicators show that indexes reflecting this information possess some significant advantages. One such index, which might be called the "duration-rate" index, gives equal weight to the duration of run and the rate of change in the indicators. This "duration-rate" index shows considerably less erratic movement than either the "percentage expanding" or the "duration of run," and it does not lag behind the "percentage expanding" index at cyclical turning points.

The ability of the "duration-rate" index to bring out cyclical trends more clearly without loss of sensitivity rests on the fact that shifts in the direction of movement of the indicators are usually preceded by a slackening in the rate of change in the indicators.

The postwar performance of the National Bureau indicators would seem to demonstrate their usefulness as a tool in the forecaster's lamentably inadequate kit. There is much room for further development of this type of approach. In particular there is need for: (1) a broadening of the coverage of the leading indicators and (2) the establishment of some theoretical linkages between movements of the indicators and movements of the aggregates to be forecast.

Field Experiments. DAVID D. MASON, *North Carolina State College.*

A range of problems in field experiments, both of a practical nature and theoretical nature, are posed and discussed. Particular attention will be given to problems in planning field experiments for technical information in agronomic, horticultural, and agricultural engineering. Presentation and illustrative examples of problems in sampling the population of conditions over which inferences are to be made, choice of treatments, plot size, and sampling schemes will be given. Some problems in the use of supplementary measurements for adjustment of data by covariance or other means will be discussed and illustrated by example.

Industrial Activity and Productivity. MILTON MOSS, *Federal Reserve Board.*

Averages of only year-to-year changes may at times obscure the underlying relations between changes in output, capital investments, labor productivity, wages, and prices. To maintain a monthly output index, manhours have served as the best alternative to measurement of areas that are particularly difficult to measure, where the production process is long and products are continually changing in complexity and are particularly difficult to standardize. Manhours and output are shown to have very similar monthly movements for the 110 series in the Federal Reserve production index based on physical product data. Also, annual changes in output per manhour have been broadly similar for the output series based on monthly product data and the output series based on adjusted manhours. The Federal Reserve is undertaking to obtain monthly data on use of electric power by type of industry group. Such data should help improve the accuracy of monthly production estimates in those areas where adequate physical quantity data are lacking and where output determinations are inevitably difficult, particularly in the metal fabricating industries. The making available of monthly power figures for important industry categories should go far in improving our understanding of relations between electricity use on the one hand, and output, manhours and other inputs, on the other hand.

Sample Surveys as a Tool for Life Insurance Research. EVA MUELLER, *University of Michigan.*

Among the research tools at the disposal of students of the life insurance market, the two most widely used are analyses of policies and sample surveys of the population. The great advantage of an analysis of policies is that one starts with company records and therefore has complete and accurate knowledge of the characteristics of the policy. The chief limitation of an analysis of policies is that it provides little information about the characteristics of owners and buyers and no information at all about non-owners and non-buyers. Hence it is argued that for most research concerned with the dynamics of the life insurance market, or with life insurance purchases as an aspect of consumer behavior, personal interview surveys are the more suitable research tool.

The validity of personal interview surveys regarding life insurance depends on sample design and sample size as well as on interviewing and questionnaire design. Insurance questions must be worded in terms which are meaningful to respondents. And the questions should not go beyond people's knowledge about life insurance in general and their own policies in particular. The potentialities and limitations of life insurance surveys are illustrated by discussing the methods of a survey of Life Insurance Ownership conducted in 1955 by the Survey Research Center of the University of Michigan for the Institute of Life Insurance.

Evaluating Case History Data at Time of Admission to a Juvenile Training Institution. PAUL MUELLER, *Department of Institutions, Olympia, Washington.*

For treatment assignment and assessment purposes a behavioral classificatory technique was developed for use with institutionalized juvenile delinquents' case histories. Narrative definitions, augmented by specific examples, define polar traits forming six behavioral dimensions: Aggression, Dominance, Non-Conformity, Gang Behavior, Anxiety, and Withdrawal from People. Dimensional reliabilities for independent ratings on seven-point scales range from .56 to .88. Three types: Socialized, Unsocialized Aggressive, and Emotionally-Disturbed Delinquents, are defined by "Ideal Type" behavioral profiles, wherein each type is assigned an "ideal" rating on each scale. Behavioral ratings for thirty delinquent boys were transformed into slope indices, i.e., comparisons of gradients between all points, and each boy classified as the delinquent type with whose "ideal slope index" his slope index had the greatest number of agreements. Use of this technique on the ratings of independent judges resulted in identical classifications in 83.3 per cent of the cases.

Further sub-classification, by means of a Personal Control score, based on ratings of Aggression, Dominance, and Non-Conformity, discriminated on a group basis between boys assigned to maximum and minimum security treatment facilities and boys who receive more or fewer severe disciplinary actions. Combined Slope Index and Personal Control score classification showed that minimum security assignments were clearly related to specific delinquent sub-types. Larger sample studies would determine reproducability of findings and rates of success for different types within varying treatment facilities.

Implications of Pension Fund Accumulations. VITO NATELLA, *Securities and Exchange Commission.*

This paper presents new projections of the financial aspects of private pension funds based on trends and developments of the last few years. In 1956 contributions to private non-insured pension funds amounted to $2.3 billion. It appears likely that contributions in 1960 and 1965 will amount to $3.2 billion and $4.4 billion, respectively. The projections imply that total assets of non-insured pension funds, which amounted to $16.6 billion at the end of last year, might increase to $29 billion by the end of 1960 and $52 billion by 1965. At the end of 1956, 52 per cent of the assets of non-insured pension funds were invested in corporate bonds, 26 per cent in common and preferred stock, and about 14 per cent in U. S. Government securities.

For the period 1951–1956 as a whole, private non-insured pension funds accounted for 17 per cent of individuals' saving. Non-insured pension funds provided about one-fourth of the security financing required by corporations in 1956. They accounted for over one-third of the new capital provided by all financial intermediaries for corporate expansion. It appears likely that the pension fund share in new financing will become even more important in the future. It does not appear from present indications that pension fund accumulations will present an investment problem in the foreseeable future under conditions of dynamic economic activity and a high level of employment.

Summary on Housing. ROBINSON NEWCOMB, *Robinson Newcomb Associates.*

The paper points out that for many purposes careful local market analyses are more important than national market analyses. The overall market may be up, while many local markets are down. Therefore even companies and mortgages interested in national markets must depend in the last analysis on local studies. Work is progressing in many centers, as Miami, Los Angeles, San Francisco, Denver, Seattle, etc. These studies provide continuous current information on production, price, and sale of new, and the price and sale of old houses, by sections of the city. In addition they provide data on vacancies, employment, income, population changes, etc., which enable builders, dealers, and mortgagees to keep unusually well informed of changing trends. Unfortunately not enough of this is being done in enough cities.

The market for the country as a whole varies in the short run with changes in 8 major phenomena: marriages, aging of the population as a whole, deaths and divorces, income and its distribution, per cent of income going to housing and commuting costs, style, quality and quantity of housing offered, price and terms, and interest rates.

Internal migration of population by itself is not listed as a major factor in the housing market, as it is not very important to either short- or long-term markets as a whole, as few areas lose population as a result of migration. Migration influences the location of the market more than its quantity.

These 8 factors may be grouped into four categories: (1) Reshuffling of the existing population. For instance, at some periods a high proportion of adults have taken residential quarters. From 1950–1955 there was a net growth of nearly 2 million new households other than husband-wife households, as compared with a net growth of just over 2 million in the number of married couple households. The ratio of growth in other than husband-wife to

husband-wife households was 3.5 times the ratio for the period 1947–1950. (2) Changing social patterns as in incomes, and in the use of these incomes. Autos were more important than housing in the 20's; housing is more important than autos today. (3) The housing and its environment offered the market. Two-bedroom houses were not popular in 1950. One-bathroom houses are not popular in 1956. New houses must keep abreast of changing tastes. (4) The price of the house. Housing prices rose precipitously from 1953–1955. Interest rates rose too. The size of the down payment, the length of the mortgage, and other terms were modified. A close watch must be kept on these four major factors, and on the influences behind these factors, such as the attitude of government officials, of mortgagees, of builders, and of families. If these attitudes can be forecast, this effect can be approximated, with luck, and usable prognostications made of the volume of home building to be anticipated.

A Program for Development of Crop and Livestock Estimates. STERLING R. NEWELL, *U. S. Department of Agriculture.*

The Agricultural Estimates Division, AMS, is a major source of agricultural statistics. The Division has recognized the need for improvement in accuracy as well as scope of its statistics in order to serve our growing economy properly. For the past four years, a research program has been conducted to investigate possible improvements in the accuracy of estimates and forecasts by probability sampling and objective measurements. Results indicate that this type of approach, used in conjunction with current procedures, would provide sizable improvements in the statistics produced.

This research by the Division has laid the foundation for a recently developed long range plan for strengthening its statistics. When fully implemented, this plan calls for enumerative surveys made in June, October, and December of each year to estimate acreages, livestock numbers, and farm labor. About 15,000 sample segments would be enumerated so that sampling errors for major crops and other items would be two per cent or less. Yield surveys in August, September, and October, and monthly surveys of prices would be made on subsamples from the June Survey. This plan would require the present budget of the Division to be almost doubled. However, the improvement in services rendered should more than justify the increased cost.

Labor Market Experience After Exhaustion of Unemployment Benefits. LEO M. ORWICZ, *Bureau of Employment Security, United States Department of Labor.*

Some 19 State Employment Security Agencies throughout the country have participated with the Bureau of Employment Security in the conduct of studies examining the personal characteristics and labor market experience of claimants who have exhausted their unemployment benefit rights. How long do such claimants remain unemployed after exhaustion? Do large proportions of them leave the labor force within two months—four months—after receiving final payment? What relationship, if any, exists between age, sex, industry or other characteristics and reemployment experience? The answers to these and similar questions are sought through these studies. Three aspects of the project will be covered by the paper: (a) the basic objectives of this type of research, (b) description of the methodology, including improvements over earlier techniques as well as limitations, and (c) a report on the study findings and their significance.

Ranking in Triple Comparisons. ROBERT NIXON PENDERGRASS, *Radford College.*

General extensions of the probability model for paired comparisons, which was developed by R. A. Bradley and M. E. Terry, are considered. The mathematical model recommended for triple comparisons is based upon the assumption of the existence of true treatment ratings, or parameters. A general class of tests of the null hypothesis that treatment ratings are equal is discussed, and maximum-likelihood estimators of the treatment parameters are obtained. These estimators are used in the evaluation of test statistics that are derived as functions of the likelihood ratio. The null hypothesis that treatment ratings are equal is tested against two special alternatives. One alternative hypothesis makes no assumptions of equality of treatment ratings. The other alternative hypothesis makes the assumption that there are only two groups of treatments and within group treatments do not differ in ratings but the two groups themselves may have different ratings. The asymptotic distributions of test statistics are discussed and approximate tests by means of the chi-square distribution are recommended. When compared with the analysis of variance, the relative efficiency of the method of triple comparisons is approximately 72%. When compared with paired comparisons, the relative efficiency of triple comparisons is 150%.

The School Leaver in the World of Work. MARGARET L. PLUNKETT, *Bureau of Labor Statistics.*

The Department of Labor, Bureau of Labor Statistics, has completed a survey of youth work experience in Clarksburg, West Virginia (Harrison County), and has studies in progress in five other widely separated areas most of which have experienced labor surpluses in recent years. Youth as defined in these studies are secondary school graduates who do not go on to college and students who drop out of secondary school before graduating. The purpose of these surveys is to find out what adjustments these young people make to the labor force and the relationship, if any, to school background. The method used in Harrison County was to get from the school records, basic data about all school leavers from September 1951 to May 1955, such as sex, age at leaving school, I.Q.'s, highest grade completed, reason for leaving, and number of vocational courses completed. From this universe, a 25 per cent sample was drawn for personal interview. From each school leaver who was still in the county in the summer of 1956 the interviewer obtained information on his reasons for leaving school, post-school training, if any, his employment and unemployment experience, and the type of jobs obtained. It was hoped that these data would throw light on why the unemployment rate of young people is so much higher than that of older age groups. If the school leaver himself was not still in the county, briefer interviews were held with a family member or a neighbor.

By the summer of 1956, nearly half of all the school leavers had left the county. Forty-four per cent of all the male school leavers were in military service at that date. Of all those who remained in the county (both boys and girls), and who regarded themselves as in the labor force, 95 per cent had found regular jobs at some time, after

leaving school, but as of the summer of 1956, 13 per cent were unemployed. Graduates started to look for work sooner than dropouts and a much higher proportion of them found jobs within a week after starting to look. The boy graduates were most likely to get their first job in some type of semi-skilled work and the boy dropouts in unskilled work. Girl graduates were most likely to become saleselerks or typists and stenographers right after leaving school, while the girl dropouts found first jobs most often as waitresses.

There was a greater difference between earnings of boys and girls in the summer of 1956 than there was between graduates and dropouts. A much higher proportion of boys than of girls earned $70 or more a week, and boy dropouts earned more than girl graduates. The girl dropouts had the most irregular employment pattern of all, and almost two-thirds of them were earning less than $30 per week in the summer of 1956.

Methods of Estimating Retail and Consumer Inventories. H. R. POWELL, *Jewel Tea Co.*

The food chain industry is rather unique among major American industries to the extent that consumer inventories of foodstuffs, both perishables and non-perishables are virtually non-existent, at least from the viewpoint of the chain operators. The food chains themselves perform the stock-carrying function and the consumer makes purchases as needed from day to day. On the whole, food chain inventories turn very rapidly although the turnover rate of 15 to 25 times each year hides the fact that inventories of perishables such as meats and produce, turn much more rapidly than do the seasonal packs of canned fruits and vegetables, for example. It is the latter category and in other dry grocery lines where some simple statistical measurements have significance. While the level of inventories is ultimately a function of sales, a controlling factor very frequently is a fixed inventory investment, expressed in the form of week's supply. Our investigations point to the conclusion that overall investment objectives of this form give rise to purchase limits and restrictions which might be inconsistent with what could be termed "economic quantities." That is, a given investment objective could result in either overstocking or understocking in specific lines or items if a simple method of interrelating the variables of merchandise flow, buying cost and carrying cost is not available. Such a method is feasible and in practice yields significant cash savings and often results in reduced inventory investment.

Seasonality in the Consumer Price Index. H. E. RILEY, *Division of Prices and Cost of Living, Bureau of Labor Statistics.*

The Bureau of Labor Statistics Consumer Price Index is subject to relatively little seasonal variation. This arises from the fact that several of the important components of the market basket, such as rent, property taxes, and transit fares, do not fluctuate in any periodic fashion, and there are broadly counterbalancing seasonal movements in other groups. The seasonal index for food reaches a peak in July, from a low point in February and March. On the other hand, apparel, fuels, house furnishings and automobiles tend to reach their highest points in the early and late months of the year, and are at low levels during the summer months. An important and as yet unresolved problem concerning the Consumer Price Index is that of adjusting the index for seasonal variations in consumption. The use of varying monthly weights provides a satisfactory index for year-to-year comparisons, but violates the essential principle of the fixed weight formula when used for month-to-month comparisons. A possible solution of this problem involves a combination of the traditional formula for yearly comparisons and an index for monthly comparisons derived by dividing a value index by a quantity index.

New Statistics on Companies and on Concentration in Manufacturing from the 1954 Census. IRVING ROTTENBERG, *Bureau of the Census.*

This paper describes briefly the recently published statistics from the 1954 Census of Manufactures on concentration, presents some additional data on concentration derived from the 1954 Census returns, makes some computations from the published figures on concentration, and describes the forthcoming data of the Census Bureau's enterprise-establishment tabulations.

Some of the new information presented is as follows: 1. Herfindahl's summary index of concentration has been computed for a sample of 48 industries. This index takes into account the size of each company individually. 2. Treating two years as a single time unit, four company concentration ratios have been computed for 20 industries. 3. Company and plant concentration ratios have been computed for the 48 industry sample, in most instances for 20 companies and 20 plants. 4. Data on imports and concentration are shown for the 43 four-digit product groups where imports are 10 per cent or more of the shipments of domestic producers. 5. The 1947–1954 change in concentration for 375 industries is shown with averages compiled by weighting the concentration ratio in each industry in four different ways: by 1947 employment, 1954 employment, 1947 value added by manufacture, and 1954 value added. All four of these computations showed slight increases in concentration—about one percentage point.

Sequential Distribution-Free Tolerance Regions. SAM C. SAUNDERS, *Boeing Airplane Company.*

The stochastic process $X_\omega = (X_1, \ldots, X_n, \ldots)$ defined on the probability space $(\Omega_\omega, A_\omega, P_\omega)$ is assumed to be generated by countably many independent replications of a random variable X on (Ω, A, P) when $(\Omega, <)$ is a linearly ordered set and P is one of a class of measures that assign measure zero to the set of equivalence classes on Ω induced by $<$.

Let $\{\lambda_n\}$ be a sequence of non-empty sets such that

1. λ_n is a subset of the first $(n+1)$ whole numbers such that the number of elements of $\bar{\lambda}_n = \{1, 2, \ldots, n+1\} - \lambda_n$ is always the same, say η, for $n \geq \eta$.
2. The set valued function $\overline{D}$ defined by

$$\overline{D}(n, x_\omega) = \sup_{j \in \lambda_n} \{\omega \in \Omega : x_j^n < \omega < x_j^n\},$$

where x_j^n is the jth ordered observation of $(x_1, \ldots, x_n)$ determined by $<$, satisfying the requirement that $\overline{D}(n, \cdot) \subset \overline{D}(n+1, \cdot)$ for $n = 1, 2, \cdots$.

Let $\{a_n\}$ be a sequence of non-negative integers such that $a_n = 0$ implies $a_{n+1} = 0$. We define the stopping event B_{n+a_n} by

$$B_{n+a_n} = [x_\omega \in \Omega_\omega : \overline{D}(n, x_\omega) = \overline{D}(n + a_n, x_\omega)].$$

Let

$$S_m = \begin{cases} B_{n+a_n} \cap \bigcap_{i=\eta}^{n-1} B_{j+a_j} & \text{if } m = n + a_n \text{ for some } n > \eta \\ \phi & \text{otherwise} \end{cases}$$

then $\{S_m\}$ is a sequential sampling plan determined by a decision rule $r = (\{\lambda_n\}, \{a_n\})$. If $\{a_n\}$ is such that $\lim_{n\to\infty} a_n/n = 0$ then $P_\omega\{\sup_m S_m\} = 1$. Then N is the random sample size defined by $N(x_\omega) = m$ *iff* $x_\omega = S_m$ and the random set function $D(N, X_\omega)$ is a tolerance region which is distribution-free in the sense that the random variable $Q(N, X_\omega)$, called the coverage and defined by

$$Q(n, x_\omega) = P_\omega[D(n, x_\omega)] \quad \textit{iff } x_\omega \in S_n,$$

has a c.d.f. which is independent of the choice of P for any r. Several different procedures are easily obtained and studied (two of which have been discussed earlier in the literature by Wilks and Jarina) by the general method outlined.

A Price Index for Custom Made Products, LLOYD SAVILLE, *Duke University*

Goods produced to the specification of the individual buyer present special problems in the construction of index numbers of prices. Because items are made to order, each new sale usually represents a distinct product and so, a new price. As a result of this, an individual establishment may quote thousands of different prices for thousands of identifiable products. And because there are frequently large numbers of firms in industries producing parts to order, millions of separate prices are charged for millions of unique products. Under these conditions of profusion and change, a sample almost inevitably is both excessively small and conservatively biased. Thus the resulting index is frequently less than adequate.

Some of these primary problems of scope and emphasis are overcome in the proposed index. By using a generalized version of the pricing formulas employed by an industry, average prices can be computed from reliable indexes of factor prices. Prices derived in this fashion may be a more accurate indication of the movements of average prices in certain industries than those achieved by traditional sampling techniques.

APHA Housing Surveys and Rehabilitation Costs. ALBERT H. SCHAAF, *University of California, Berkeley.*

Surveys based on the American Public Health Association's method of appraising housing quality are widely used to guide local housing programs. With rehabilitation assuming a key role in many Urban Renewal projects, it is important that an APHA survey show its possibilities. This study tests the significance of the relationship between APHA dwelling unit penalty points and rehabilitation costs. A correlation analysis of cost and point data for 47 structures located in Oakland, California, shows that costs and points are significantly interrelated and that a useful cost-estimating regression formula can be developed by compiling cost data for a small sample of structures drawn from those included in an APHA survey. The methodological technique of making such an analysis, however, requires a consideration of several factors: 1. A uniform definition of rehabilitation is necessary and should be most meaningful and realistic when applied to the structures and the area. 2. Since the APHA index does not specifically reflect relative costs, weighting of individual items in the index may be necessary. 3. Measurement problems are presented because structures typically vary in number of dwelling units and in size and because the penalty points refer to dwelling units while the cost estimates refer to structures.

The Implications of the Changing Ownership of Federal Securities. IRA O. SCOTT, JR., *University of Minnesota.*

During the postwar decade, the shares of the Federal debt held by commercial banks, mutual savings banks, and insurance companies have declined. Those of the U. S. Government investment accounts, state and local governments, and miscellaneous investors have increased. Corporations, individuals, and the Federal Reserve banks have maintained about the same relative position. Unlike the other investor groups, the U. S. Government investment accounts are, for the most part, restricted by law to Federal securities. Assuming no increase in the Federal debt, the U.S. Government investment accounts will present an inelastic demand for a fixed supply of Federal securities.

Eight out of fifty-two trust funds and other accounts handled by the Treasury presently hold the bulk of Federal securities owned by these accounts. Five of these funds, The Federal Deposit Insurance Corporation, the Civil Service Retirement and Disability Fund, the Federal Old-Age and Survivors Insurance Trust Fund and Disability Insurance Trust Fund, the Railroad Retirement Account, and the Unemployment Trust Fund, will expand an estimated $3 billion, $6 billion, $35 billion, $2 billion, and $2 billion, respectively, or a total of $48 billion, by 1975. The projected expansion of the trust accounts will absorb over 60 per cent of the outstanding marketable debt. This development will greatly improve the Government bond market and facilitate the Treasury's task of lengthening the term structure of the debt.

Drop-Outs in High-School in Large Cities. DAVID SEGEL, *Office of Education.*

A cooperative study of drop-outs in large cities—over 200,000 population—was carried out over the four years required from the entrance of the freshmen in high school to their graduation. One objective was to create a new method for pupil accounting in secondary schools and to explore its effectiveness in operation. A second objective consisted of noting significant facts uncovered through this superior method of accounting and suggesting possible improvements in the school program.

The objectives of the study were attained. The superiority of the accounting procedures over the older survival rate method was clear and several facets of the data were of considerable importance to school programs. Among

the facts which seemed important were (1) school leaving was concentrated around the tenth grade and (2) a large number of students transferred among the schools within a city and also among cities.

Certain suggestions for improvement within the schools seemed to grow out of the findings of the study. Among these suggestions were:

1. The curriculum of the high school needs to offer curriculum sequences adapted to the characteristics of the students beginning with the 9th or 10th grade. The present curriculum seems better adapted to girls than to boys.
2. There is need for a concentrated guidance attack at the 9th grade level in giving students a view of the educational and occupational world, and help in the analysis of each individual in relation to his potentialities so that assuming the school offers curriculums suited to these potentialities he can be aided in choosing and entering them. Through this process more students will stay in school until graduation and become assets to our country as well as to themselves.
3. The great amount of transfer between schools indicates that, at least in each city, the high schools—except for special trade schools—should all offer similar curricular sequences and subject offerings. This would alleviate for students the difficulty of adjusting to a new curriculum when transferring. It is possible that the confusion resulting in trying to make an adjustment to a new school may be one of the major factors in the school leaving problem.

Russian Population Censuses: Methods and Concepts. GALENA V. SELEGEN and MICHAEL K. ROOP.

A new population census has been decreed for January 1959 in the Soviet Union. Preliminary plans for the census were drawn up by leading statisticians in various conferences during July 1957, and on August 1, 1957 these plans were subjected to a test run in 7 selected districts of the USSR. One of the first clues which indicated that the Soviet government was preparing for a new population census was the publication in 1953 of a 150-page manual concerning the last (1939) all-union census. The methods and concepts employed in the new census will, of course, be greatly influenced by Russia and the Soviet Union's past experience in census-taking and many of the problems which will be encountered can be anticipated by a consideration of the leading censuses of the past.

The forthcoming census will be the fourth to cover the entire territory of the USSR (earlier comprehensive censuses were taken in 1920, 1926, and 1939), although there was one general census of the inhabitants of the former Russian Empire, in 1897. During the last half century the methods and concepts of these censuses were subject to considerable change, in part due to the progress of science but still more to exigencies of Soviet economic planning and propaganda objectives—the latter two dominant since the nineteen thirties. In the 1939 census as compared with the census of 1926, 4 new questions were introduced to improve the accuracy of the count, while 4 questions were dropped and the meaning of several questions changed by definition, in great part to make the census correspond to an explicitly stated notion that the census should be conducted in a manner which would support the regime's political objectives. Further, the detailed results of the 1939 census, unlike the 1926 census published in 50-odd volumes, was considered a state secret, only bare summaries of findings being made public. There is no doubt as to the high level of accuracy on which the prospective 1959 census will be conducted and tabulated as regards the needs of the Soviet administration and planning agencies. However, it remains to be seen the extent to which the results of the census will be objectively reported or made public in full detail, and the outlook at present in this respect is not promising.

Seasonal Adjustments of Economic Indicators—A Progress Report. JULIUS SHISKIN, *Bureau of the Census.*

This paper reports on the experience accumulated during the past year in a research program on ways of dealing with certain problems that arise in making seasonal adjustments of economic series.

1. An analysis of a sample of about 150 economic series shows that for the post-World War II period seasonal movements dominate the month-to-month cyclical and irregular movements in economic series. Thus, the average month-to-month amplitude of the seasonal fluctuations exceeds that of the cycle-trend factor in 78 per cent of the series, that of the irregular factor in 65 per cent of the series, and the cycle-trend and irregular factors in combination in 45 per cent of the series. This means that it is usually easy to isolate the seasonal factor but not the cyclical factor.

2. A test of the Univac method of decomposition was made by taking irregular, cyclical, and seasonal components computed for real economic series, scrambling the components of different series, decomposing the synthetic aggregates by the electronic computer, and then comparing the components put into the synthetic series with those yielded in the Univac decomposition. While the results generally appear on the basis of inspection to be satisfactory, they also show that the magnitude of the largest of the three factors is always reduced, with offsetting increases in one or both of the other factors. Where two of the factors are large relative to the third, the smallest factor can be significantly affected. The number of national series in which the smallest factor will be significantly affected is, however, quite small, probably less than five per cent. The proportion of individual company or local area series so affected may be much larger.

The synthetic series were also used to test the contribution of the second round of computations, based on the weighted fifteen-term moving average, in Method II. The results of these tests are clear: this iterative technique substantially improves the seasonal adjustments.

3. A test of alternative confidence limits for identifying extreme seasonal-irregular ratios shows that the narrower the limits, the smoother the seasonal adjustment factor curves. We have not yet been able to establish objective criteria for determining the appropriate amount of smoothness to be secured by this step. After a study of the charts and tables constructed for this test, we believe intuitively that it is desirable to shift from control limits of two to one and one-half standard errors, and we plan to do so soon.

4. Seasonal patterns of aggregate series will ordinarily change because of changes in the relative importance of the components, and the impact of this factor on the seasonal pattern of total U. S. exports is shown. No measures of the general importance of this factor are, however, given.

5. A test was made of alternative ways of computing seasonal factors for the ends of series—leveling off or allowing for a continuation of the trend. The tests compared the results obtained in both methods when only partial data are available and when full data are available. Five series, with different mixtures of irregular, cyclical, and seasonal components, were used. Both the inter-war and the post-World War II periods were covered. The test showed that the Method II technique of allowing for a continuation of trend gives better results than leveling off at the fourth from end year, for which no estimating is involved. We made similar comparisons with the third from end year even though it does include estimated ratios, in Method II. The use of the factors for this year and the Method II technique gave estimates of about the same accuracy. Whether this is a statistical accident or points to a way of improving the Method II technique, I do not yet know.

A test was also made of alternative ways of estimating the twelve seasonal factors for the year ahead. This test clearly supports the theoretical position that it is better to include a trend allowance in making such estimates.

On the basis of these tests, similar tests of other techniques of allowing for trends at the ends of series have been scheduled, and a routine has been added to the Univac program estimating seasonal factors one year ahead with a trend allowance.

6. The plans for future extensions and tests of this program are described briefly. The principal extensions we have in mind are: (1) confidence limits for judging the significance of month-to-month changes, (2) tests of significance for the $\bar{I}$, $\bar{C}$, and $\bar{S}$ measures, (3) computations of the $\bar{I}$, $\bar{C}$, and $\bar{S}$ measures for different stages of the business-cycle, (4) a measure of secular trend, and (5) a semi-log chart program. We also plan to make tests of (1) alternative weight patterns for computing the seasonal factors for the end years and for the year ahead, and (2) alternative curves for measuring the cycle-trend factor.

7. The studies made throughout the year emphasize the importance of careful professional review of the electronic computer results, particularly at the ends of series.

Measurement Problems in Collecting Supplementary Data about Deceased Persons. MONROE SIRKEN, *Public Health Service.*

The information contained on the death certificate provides the basic tabulations of mortality statistics, but frequently there is a need for supplementary mortality statistics, particularly for epidemiological studies. The main objective of this paper is to compare the quality of data on the decedent's maximum smoking rates obtained by mail survey and personal interview with close relatives. In a recent study, smoking histories for a sample of deceased persons in Pennsylvania were collected initially by a mail survey of relatives. Nonresponse interviews were conducted on a subsample of nonrespondents to the mail query, and quality check interviews were conducted on a subsample of respondents who returned the mail form. The quality check interviewer collected the data on smoking habits without reference to the results obtained by mail query which were posted at the end of the schedule. After the interview was completed, the interviewer reconciled differences between the reports obtained by mail query and personal interview with the assistance of the respondent.

The quality of the data collected by mail and personal interview is evaluated from the viewpoint of nonresponse bias and response differences. The effects of two separate conditions on response differences in the reported maximum smoking rate are noted: (1) the effects of different respondents replying to the mail query and the personal interview, and (2) the effects of memory distortion because the decedent was not smoking at the maximum rate during the 1-year period before death.

Methods of Inquiry—The Consumption Approach. ELEANOR M. SNYDER and JOHN G. MYERS, *Interdepartmental Committee on Low Incomes, New York State.*

Investigators concerned with describing the characteristics of the population with low economic status or with measuring the relative size of this group within a given population are hampered by the limitations of income size distributions which cover income received during the period of one year. In the absence of data on income received over a period of several years, for identical families, indirect methods of research must be employed. Two methods of distinguishing families with low economic status from those with temporarily low incomes are examined. The consumption expenditures of individual families are employed in both methods to identify those with "permanent" and those with "transitory" low incomes. The first method utilizes expenditure data on specific goods or categories of goods; the second method is based on a classification of families by total consumption expenditures. The data are drawn from a series of special tabulations of the 1950 Survey of Consumer Expenditures conducted by the U. S. Bureau of Labor Statistics. The first procedure was developed in a study sponsored by the Franklin D. Roosevelt Foundation and conducted by E. M. Snyder. This study was the subject of a paper "A Method of Identifying Chronic Low Income Groups from Cross-Section Data," presented to the March 1956 Conference on Research in Income and Wealth. The second procedure was suggested by William Vickery and given theoretical support by the recent work of Milton Friedman, Franco Modigliani, and Richard Brumberg. The City Workers Family Budget is used as a low income "yardstick." The results of both procedures are compared, examining the number, proportion, and characteristics of families with low economic status found by each method.

A Decision Model for Computer Utilization. MORRIS J. SOLOMON, *American Greetings Corporation.*

A general model is presented which enables a programmer to choose a procedure with the least cost, given the probability of failure, the costs of failure, and the cost of measures to control or ameliorate the failure. Two decisions which are facilitated by this model are: 1. Whether a machine produced tape should be read back before proceeding, and 2. The desirable length of machine running between recovery intervals.

On the Logic of Diffusion Indexes. ALBERT T. SOMMERS, *National Industrial Conference Board.*

This paper explores the logic of diffusion indexes in terms of the systematic uncertainties introduced in the course of the transfer of diffusion analysis from structural studies of the cycle to mechanical forecasting.

Diffusion indexes take their materials from structural studies of diffusion, but focus upon timing relationships among specific economic series. In this sense, they are an "indicator" approach to forecasting, distinguishable from such approaches popular in the past by greater sophistication in the testing of indicators, and by the multiple use of indicators, which permits gradual changes in probabilities, and opens the attractive (but partly illusory) possibility that irregularities in the behavior of indicators will tend to offset.

The insufficiencies of evidence provided by a diffusion index arise for three classes of reasons. I. Difficulties arise with respect to interpreting the movement of individual series. Even series with good conformity characteristics when viewed historically may contract, month to month, almost as many times as they expand during the expansion phase and vice versa. Smoothing devices designed to reduce the irregularities tend to erode the timing lead, and may distort the cyclical timing itself. In addition, the presence of secular trends, and their varying rate from cycle to cycle, raise uncertainties about the location of cyclical turning points or, to put it the other way around, introduce uncertainties into the implications of a turning point once it is determined. Reduction of irregular and secular components in movements can be achieved only by examining the series partly retrospectively (as in the form of centered moving averages) and/or by moving from binary decisions about turning points to probability weightings (as in duration-of-run procedures).

II. Uncertainty enters into the interpretation of the index because of its fixed-content, fixed-weight composition. The widely recognized "uniqueness" of each cycle is more characteristic of peaks and troughs, with which diffusion indexes are concerned, than of the continuous phases of expansion and contraction that lie between them. The "content"—as distinct from the "scope"—of turning points varies from cycle to cycle; the change in "scope" recorded in the index's fixed range will vary depending upon whether the initial changes appear fully within, or only partly within, or not at all within its range.

III. Efforts to achieve very broad coverage raise another set of uncertainties. The criteria by which individual indicators are judged for inclusion in diffusion indexes—degree of lead, reliability of timing, regularity of intraphase movement—appear to be inversely correlated in the performance of individual series. Performance of the collective index against one criterion can thus be improved only at some cost, in the form of reducing its performance against one or more of the other criteria. Finally, coverage of the collective index is debarred from including series whose timing characteristics are found to be "irregular," (since the turning point in such a series, assuming it can be identified, carries no systematic implication for the general cycle). Irregular timing is found among those activities which are not controllable over the course of the cycle or which do not conform because their motivation is not related to the cycle. The same reasons that forestall the use of this series in diffusion measures act to reduce uncertainties in more subjective forecasting procedures.

Cyclical Behavior of Manufactures' Inventories, 1945–1955. T. M. STANBACK, JR., *New York University.*

This study is principally an analysis of Department of Commerce monthly inventory data for industry groups along with 25 commodity series for finished goods stocks. Value data for 16 industry groups have been deflated since 1947, nine of which were deflated separately for finished goods and purchased materials.

Fluctuations in inventory investment are shown to have been a major force contributing to changes in gross national product during both post war recessions, with changes in manufacturers' inventories playing the leading role among the several categories of stocks. The use of monthly and quarterly data permitted more precise statistical description of inventory behavior than was possible in studies of the prewar period. This fact was especially important in the observation of inventory investment series in which well-developed cyclical movements of rather brief duration occurred. Typically, two such investment cycles occurred in our inventory series during each reference cycle, the second such investment cycle tending to conform closely to reference cycle turning points. Cycles in the durable manufacturers category showed both better conformity and greater amplitude than nondurables pointing up the former group as the more important source of inventory instability. Analyzed on a stage of fabrication basis inventory series showed strong cyclical patterns for purchased materials and for finished goods investment. These patterns are described in detail and their implication for revision of cycle theory discussed.

Ecology of Coronary Heart Disease: Statistical Implications. MORTON ROBINS and ALFRED M. STEINMAN, M.D. *Veterans Administration.*

The past two years have witnessed a tremendous surge of interest in the ecology of coronary heart disease. A conference on the epidemiology of arteriosclerosis and hypertension, a symposium on the results observed in epidemiological investigations of coronary heart disease, and numerous papers on the diagnostic, therapeutic, preventative and etiologic aspects of this leading cause of death—all of these recent additions to the literature have whetted the appetite of both the medical and public health professions for more definitive data. The lay public is desirous of immediate programs for prevention. However, despite the suggestive and widely publicized association between diet and coronary heart disease, a control program does not appear on the immediate horizon.

This paper considers the various ecologic factors currently believed to influence the pathogenesis of coronary heart disease. An appraisal is made of the methodological problems faced by the biometrician and epidemiologist in verifying and quantifying these hypotheses as to etiology.

Examination and assessment is made of the advantages and limitations of each of these methodological approaches: (1) demographic studies on data obtained from morbidity and mortality reports; (2) retrospective studies; and (3) prospective studies.

Conclusions, or perhaps our recommendations, will suggest the following:

(1) Studies of selected special groups believed to have an appreciable greater susceptibility to coronary artery disease, e.g. diabetics; (2) Retrospective studies of hospital and clinic populations with and without coronary artery disease; (3) Intensive, short-range, community programs for morbidity reporting of each case of coronary artery disease with specified ecologic data.

How to Communicate Statistical Findings in the Field of International Markets to Management. J. WILNER SUNDELSON, *Ford International Division, Ford Motor Company.*

Periodicity, methods and forms of presentations of economic and market materials are fully comparable with those used by domestic divisions of Ford Motor Company. Major differences with domestic reporting lies in availability of international statistical and economic data as well as in comparability of material coming from multiple markets.

International automotive, tractor, and parts and accessories data are relatively poor in quality, irregular, disorganised, and delayed. In foreign manufacturing centers only Germany approaches comparability with U. S.; all others are much worse. Basic reason is that in most manufacturing countries neither government nor trade associations collect and disseminate data as is done in the United States. Furthermore, in very few countries do leading manufacturers, exporters or importers informally exchange (within the trade) historical and current performance data.

Basically U. S. data recognizes domestic manufacture and import sources and analyzes all stages and aspects of market developments. Wide variations in foreign markets result in data which cover production, exports, imports, wholesale deliveries, and registrations. Supply cycles are much longer than those of U. S. domestic market and reporting periods are different adding to complexity of problem of reducing world-wide sales results to comparable basis.

Ford International Division's compromise has been to collect best available data from all sources and to reduce these to some unified picture. Best statistical market data represent production at manufactured sources, shipments, and wholesale deliveries by assembly plants, to dealers and distributors in export markets. Adjustments are made for transshipments wherever possible.

Forecasting of market volumes is an important aspect due to time lags in implementation of supply, product, and facility programs. Practice has been to have staffs in large manufacturing sources develop market trend data and forecasts for their own domestic markets. Ford International, Manufacturing Assembly and Sales locations, cooperate in the development of forecasts for all other important markets. Short term estimates related to early budgeting procedures. Long term forecasts organised around concept of Standard Volumes. Following formal acceptance of Standard Volumes these are used for profit, supply, product and facility and all other forms of planning.

Ford International Division now concentrates collection of historical and current sales information and short term forecasting in Marketing Office. Economic and long term forecasting including Standard Volume studies are responsibilities of Operations Planning Department in Facilities and Operations Planning Office.

Factors Underlying Productivity: Some Empirical Observations. NESTOR E. TERLECKYJ, *National Bureau of Economic Research, Inc.*

The purpose of the paper is to inquire into the possibility of finding statistically significant relationships between the available measures of productivity change and several explanatory factors. Productivity change is defined in terms of total factor productivity. The measures of productivity change are available for several manufacturing industries and for several periods. These unpublished data were made available by Dr. John W. Kendrick of the National Bureau of Economic Research. Matching measures were compiled for the following explanatory variables: rate of growth of the industry, amplitude and frequency of the specific cycle, capital-labor ratio, firm and plant size, intensity of research activity, prevalence of professional technical personnel, rate of entry of new firms, and index of concentration. On the basis of simple rank correlations there are indications that growth of the industry, its cyclical stability, and intensity of research may contribute to progress in productivity. Further statistical processing of the available material, particularly by multiple regressions, is planned.

Some Estimates of Tolerance Limits from Censored Data. EARL A. THOMAS, *Avco Research & Advanced Development.*

The problem of estimating statistical tolerance limits on the performance or reliability of equipment has received renewed emphasis in recent years. However for physical, economical or accidental reasons data from the tails of the distribution are often missing or suspect. This situation results in data from a "censored sample." Statistical Tolerance Limits are defined as by Wilks, Censored and Truncated Samples are classified as by Hald, Gupta et al. The simple (in application) distribution free techniques of R. B. Murphy are described. The work of Epstein and Sobel and that of Deming and Votaw have been utilized to estimate tolerance limits from censored samples of two types of exponential distributions. The results of empirical experiments using data from the normal distribution are reported. It is apparent that combining the estimates from order statistics obtained by Sarhan and Greenberg with the values of Bowker yields workable results for sample sizes up to size twenty. Maximum likelihood estimates derived by Gupta can be combined satisfactorily with Bowker's values for sample sizes of approximately 100 and greater. No adequate tables exist for intermediate size samples.

Approximately Distribution-Free $\bar{x} \pm ks$ Type Tolerance Intervals. JOHN E. WALSH, *Lockheed Aircraft Corporation.*

This paper contains an investigation of distribution-free applications of $\bar{x} \pm ks$ type tolerance intervals for a moderately general class of continuous populations. This class consists of the populations whose distribution can be adequately represented by a generalization of the first seven terms of its Edgeworth expansion. To first order terms with respect to sample size n, the $\bar{x} \pm ks_j$ tolerance interval is found to depend only on s_j. Here $\bar{x}$ is computed from all n sample values but s_j is based on a subsample whose size is of order $n^{1/2}$. Using this property and dividing the sample into $r = \phi(n^{1/2})$ independent subsamples, of approximately the same size, the first order distribution of the $\bar{x} \pm k \max(s_1, \ldots, s_r)$ tolerance interval is obtained. On the basis of its first order distribution, lower bounds are stated for upper probability points of the distribution of this tolerance interval. For sample sizes of 25 and larger, these bounds seem to be valid for a large subclass of the class of continuous populations initially considered. This

subclass appears capable of approximating most continuous populations of practical interest. By suitable selection of k and n, a large number of different tolerance intervals can be obtained. For many cases of practical interest, the sample size required for a specified tolerance interval of this type is noticeably less than that for the corresponding distribution-free tolerance interval based on the largest and smallest sample values.

Characteristics of Life Insurance Sales and Buyers Which Influence Persistency in Premium Payment. ALFRED G. WHITNEY, *Life Insurance Agency Management Association.*

This paper is a report of the development of methods for predicting life insurance premium-paying behavior. Since the termination of a life insurance policy represents a failure in a plan for security, life insurance companies are eager to find factors associated with persistency in premium payment which can be used in training their agents in the methods of identifying potentially persistent clients. The data for this study came from a sample of 24,817 policies issued in May 1949. Two years after issue a check was made to determine whether or not each policy was still in force. An analysis was then made of the information about the policy and policyholder which is available at the time of sale and persistency rates were calculated for a list of variables. Further analysis of the relationship with persistency for the most predictive variables and of their interrelationships led to the development of a Persistency Rater. Copies of this Rater which now has extensive use in the life insurance industry were given out at the time of presentation of the paper.

Unemployment Effects of a Major Plant Shutdown in a Depressed Area. RICHARD C. WILCOCK, *University of Illinois.*

What happens when the dominant industrial plant shuts down in a labor surplus area? This situation was studied in a southern Illinois community by the University of Illinois for the Bureau of Labor Statistics. The paper presents some findings on the unemployment effects of the shutdown.

Prior to the shutdown, the community had had chronic unemployment because of employment declines in coal mining and agriculture. The shutdown plant had been the dominant factory employer and its closing accentuated the unemployment problem. Although worker out-migration was extensive, continued employment declines in the area and reluctance of workers to leave the community kept the unemployment rate high. In the two years after the shutdown, the unemployment rate fell from 16.3 to 11.1 per cent. The labor force, however, became 13 per cent smaller.

In the two years following the shutdown, the average worker from the closed plant experienced more than 6 months unemployment, but workers without farm land, older workers, and those with less education had longer unemployment than others. At the time of the study, a third of the workers had jobs in other areas; ¼ were unemployed, underemployed, or had left the labor force; and ¼ had full-time jobs within the area. These and other data illustrate a very strong home-town attachment and help explain the persistence of chronic unemployment.

Subject Matter Covered in Engineering Statistics at Virginia Polytechnic Institute. R. LOWELL WINE, *Hollins College.*

The two quarter course (about 60 lecture hours) is open to seniors and graduate students in engineering. No prior knowledge of statistics is assumed. The object in offering such a course is to acquaint the student with basic statistical terms and concepts. It is hoped that on completing the course students will be able to apply elementary statistics and know when to call on a professional statistician for advice when faced with the more complicated statistical problems.

Topics discussed include measures of central tendency, measures of dispersion, population and sample, unbiasedness and efficiency, the normal distribution and the central limit theorem, statistical inference with emphasis on tests of hypotheses, analysis of variance methods applied to one-way and two-way classifications, design of experiments along with orthogonal comparisons and multiple comparisons tests, regression and correlation, one-way covariance analysis, goodness of fit tests, control charts and single sampling plans. Approximately half of the time is devoted to tests of hypotheses, analysis of variance, and regression analysis.

A New Approach to the Estimation of Inventories. RALPH S. WOODRUFF, *Bureau of the Census.*

It has been very difficult to obtain monthly estimates of retail inventories because most businesses do not take inventories monthly and it is an unreasonable request to ask them to count their complete inventories solely for statistical purposes. The proposal is to obtain estimates of total inventories in all retail establishments by drawing a random sample of establishments and within each sampled establishment drawing a small random sample of specific commodities to be actually counted and priced. In this way the workload for any given establishment can be reduced to the point where either the respondent or an enumerator can usually make the requested counts in a matter of minutes. This approach if successful will yield, in addition to estimates of total dollar volume of inventories of all commodities, (1) estimates by type of commodity and (2) estimates of physical quantities either directly or in index form. This approach has the disadvantages (1) that the estimates are only as complete as the basic list of commodities used in the commodity sampling and (2) that additional variance due to the sampling of commodities within establishments is introduced.

The proposed system has not yet been put into overall operation at the Bureau of the Census. However, a list of commodities used in retail has been developed for this purpose and experimentation on a limited scale with obtaining actual counts of specific commodities has been undertaken. In addition, the principle has been used in a survey of food stocks made by the Bureau of the Census for the Federal Civil Defense Agency.

BOOK REVIEWS

Nonparametric and Shortcut Statistics. *Merle W. Tate and Richard C. Clelland.* Danville, Illinois: Interstate Printers and Publishers, Inc., 1957. Pp. ix, 171. $3.95. Paper.

WILLIAM H. KRUSKAL, *University of Chicago**

THIS book is primarily an elementary exposition of applied nonparametric analysis for research workers in the social, biological, and medical sciences. The first third of the text proper is a general introduction to statistical inference. Some short-cut parametric techniques, mostly involving the use of sample ranges, are briefly discussed. There are 28 pages of tables and graphs. Numerous problems with real data are set, and answers or discussions for them are collected at the end of the book. Printing appears to have been done directly from typescript.

The book, in my opinion, does not adequately fulfill its purposes, and the world still awaits an acceptable introductory textbook for applied nonparametric analysis. There are several general reasons for this critical evaluation. First, the book shares with its competitor by Signey Siegel [8] an unfortunately high density of errors and ambiguities.[1] Second, examples are discussed in a formal, shallow manner so that the reader keeps asking "But what are you really trying to find out? What about the sampling method? What are possible biases? What are the alternative hypotheses?" Third, several important references are not listed, nor are their contents reflected in the text.

On the positive side, the authors are to be congratulated for paying some attention to techniques other than tests of no-difference hypotheses, especially confidence intervals, and for discussing not only advantages but disadvantages of nonparametric methods. Unfortunately, they do not go nearly far enough in their discussion of confidence intervals, and their presentation of the limitations of nonparametric procedures is incomplete and sometimes in error.

Book reviews in this and other statistical journals complain again and again of a high density of errors. Does this complaint stem merely from the querulous pedantry of overmeticulous reviewers? I think not, but rather than standards of accuracy and clarity for statistical texts are unhappily low, perhaps lower than those for texts in other fields. The only long-run remedy for the problem will be a refusal of the statistical public to recommend and buy inferior texts.

It might be said that the responsibility for accuracy and clarity is that of the author. But I feel that, given the present complex relationship between publishers and the scientific community, a considerable responsibility also falls on the publisher. He is in a position to see that manuscripts are carefully read by independent competent referees, and to urge appropriate corrections upon the authors. Both in the case of Siegel's text and the present one, such a procedure would have resulted in incomparably better books, and propagation of error would have been materially reduced. Good refereeing, to be sure, must be paid for, in time and energy on the

* Several people have commented on a draft of this review, and their remarks have been helpful in revising the draft for greater clarity. The opinions expressed in the review are, of course, not necessarily shared by those who commented. Besides my colleagues at Chicago I should like to thank the following: J. Gilbert (Center for Advanced Study in the Behavioral Sciences), F. Mosteller (Harvard), I. R. Savage (Minnesota), and J. Stanley (Wisconsin). In addition, I wish to express my appreciation to M. W. Tate and R. C. Clelland for their gracious and useful comments.

[1] Siegel's book has been competently and carefully reviewed by I. R. Savage in this *Journal* [7]. Many of Savage's criticisms of Siegel's book apply as well to the book by Tate and Clelland; and Savage lists alternative expository sources.

part of the referee, and in cash on the part of the publisher. But five or six hundred dollars paid for refereeing, especially when spread over sales of, say, 4000 copies is not too onerous. I would far rather pay $4.15 for a properly refereed book than $4.00 for one that is heavy with troublesome slips, ambiguities, and downright errors.

APPENDIX

The remainder of this review lists some of the more egregious or otherwise noteworthy errors, ambiguities and omissions that I noted in reading the book, and discusses contents and tables, especially in comparison with Siegel's book.

P. vi. " . . . nonparametric test probabilities [P values] as large as about .20 tend to be associated with conventional rejection levels under more powerful tests." This startling statement appears in the preface and I find no evidence given for it there or elsewhere. It might mean that, if we consider samples that have led to conventional rejection level P values with parametric tests, then we tend to find appreciably higher P values with nonparametric tests. Or it might mean that actions taken when a parametric test gives a P value of about .05 tend to be the same as those taken when a nonparametric test gives a P value of, say, .15. In either case, I question the statement, although I do not think anyone has studied either interpretation. Whatever the motivation, nonetheless, we should be pleased that the tables given by Tate and Clelland run up to .20 significance levels.

P. 2. "Only a random sample can be considered unbiased and therefore representative . . . ." First, it is not made clear that it is the *method* of sampling that is random, unbiased, etc., not the particular sample at hand. Second, other methods of sampling may well result in unbiased and representative results, however those adjectives may reasonably be defined. For example, stratification and clustering are well-known devices of sampling that do not give random samples in the strict sense. This point is hinted at on p. 24 but not made clear.

P. 3. "Sample data . . . are of two broad kinds, *quantitative* and *qualitative*. The former . . . can be arranged in order of size. Qualitative data do not allow such an arrangement." Surely any book in part for social scientists should consider ordered categories for qualitative data.

P. 7. " . . . [sample cumulative distribution] curves do not provide any very exact information about the data." On the contrary, they provide, for random samples, *all* the information about the data. Further, two well-known nonparametric procedures, both discussed in this book, are directly based on the sample cumulatives.

P. 9. "Inferences drawn from a sample median about the population median, although not as accurate as those drawn from a sample mean, are safer and more general in the sense that they do not rest on the assumption of normality." But a sample median may be quite irrelevant to the population mean, and when it is relevant, it may provide a more accurate estimator than the sample mean.

P. 11. "It is sometimes convenient to speak about [the] modal . . . category [for qualitative data], but used this way the mode has no statistical meaning."

P. 15. "[Kendall's τ] is not as good as [Spearman's r_S] in describing relationship for it [τ] is not directly related to the product-moment coefficients." Why should direct relationship to the product-moment coefficients be a criterion?

P. 16. "τ . . . is always smaller than r_S except near unity or zero, where the two tend to agree." It is true that when τ or r_S is near ± 1, so is the other, but the remaining two parts of the statement are false or misleading as may be seen from the following example. Consider the paired rankings

1	2	3	4	5	6	7	8	9	10
9	10	1	2	3	4	5	6	7	8

for which $r_S = 1/33$ and $\tau = 13/45$. Also, the quoted statement is ambiguous about whether absolute or algebraic values are meant. Correct inequalities between r_S and τ are given, for example, by Kendall [3].

P. 24. The introduction to hypothesis testing discusses a matched pair experiment without a word about randomizing treatments.

P. 36. After some fine comments earlier about the importance of considering alternative hypotheses in framing a test, the test proposed here for a simple multinomial null hypothesis is obtained by including in the critical region sample points with lowest probabilities under the null hypothesis.

P. 42. "When we have a sample statistic, we may use it to estimate the corresponding parameter. This is known as *point estimation*." This seems to put the cart before the horse. The first problem in point estimation is to decide what population characteristics are of major interest.

P. 44. and many other places. The reader gets an impression from this book that only for underlying normal distributions are parametric methods useful.

P. 52. Concerning a χ^2 test on a trinomial population, "The fact that we cannot state a definite alternative makes it impossible to determine the power of the test. Nor is it possible to estimate [sic] confidence intervals for the [population proportions] . . . ." But approximate power functions and confidence sets in this case are, if not well known, at least known to many. Power is discussed, for example, by P. B. Patnaik [6]. Confidence intervals for a single one of the trinomial or multinomial probabilities may be found just as in the binomial case. Approximate joint confidence ellipsoids may be obtained for all the probabilities by including in the confidence region, in standard fashion, all those sets of probabilities that would, as null hypotheses, be accepted by the sample at hand.

P. 53, Example 21. This gives the post positions of the winning horses in 144 races, and the reader is asked whether it is reasonable to believe that post position doesn't affect the chance of winning. The "correct answer" at the back of the book is to apply a standard χ^2 test. Here is a perfect place to consider ordered alternatives, but not a word is said. It is of interest that Siegel discusses the same example [8, 44–6] and does mention the order question, although only in a footnote and without presenting any appropriate method.

P. 56. Brief mention is made of application of the Wilcoxon signed rank test to a single sample, but no cautionary remark appears about the special nature of the test in this case. (When it is used to test or estimate a translation parameter, symmetry must be assumed.)

P. 60. "... [a disadvantage] of the conventional χ^2 test for goodness of fit [is that] if the signs of the differences are alike in several consecutive intervals of the distribution, the test may fail to detect significant departures from the assumed curve ... nothing can be done about [this] ... except care in interpretation of the test." The disadvantage is not clearly described and one can take signs of deviations into account.

P. 63. In finding graphically the maximum vertical deviation between hypothesized and sample cumulatives it is suggested that probability paper be used that distorts the vertical scale so that the hypothesized cumulative is a straight line. This may make it rather hard to find the maximum vertical deviation.

P. 74. In discussing Fisher's exact 2×2 test (not identified by this name) no reference is made to the several available tables for it, nor is any reference made to Table J of the Tate-Clelland book itself which provides some approximate critical values for the 2×2 situation.

Chap. IV. In the discussion of chi-square tests, despite earlier statements of good intention, there is no discussion of power.

Pp. 78–82. The discussion of nonparametric linear regression considers only the Brown-Mood method and omits any mention of the Daniels method [2].

P. 84. "None [of the two sample nonparametric tests] lend themselves to quick or accurate estimation of confidence limits." The Wilcoxon tests lead to quite easy graphical methods for obtaining (not estimating) confidence limits. Details are partly discussed by Moses in [5], a reference listed by Tate and Clelland. The same point arises on p. 103.

P. 85. "The connection between the contingency test of independence and the test of sample proportions arises from the fact that the data in two samples may be classified in the 2×2 table." This is a little like saying that the connection between adults and children is that both have four limbs.

P. 90. "The [Wilcoxon two-sample] test is the nonparametric test most nearly analogous to the standard parametric t test. ..." Surely, it is not as analogous as the permutation test on the observations using the t statistic, nor as analogous as the Fisher-Yates normalized rank method. Further, any test based on permutations of a set of preassigned numbers, whether ranks, normal order statistics, or some other set, is at least as much an analog to the t test as Wilcoxon's test. A parallel comment may be made about the same remark in the paired two-sample case (p. 101).

P. 90. "It [the Wilcoxon two-sample test] assumes neither normality nor homogeneity of variance." It *does*, as usually presented, assume not only identity of variance but identity of distribution under the null hypothesis. To be sure, it may not be so very sensitive to some aspects of this assumption, but this is another matter.

P. 100. Of the sign test applied to the so-called correlated 2×2 situation: "It would be very difficult to study the power of the test and, practically speaking, idle to do so." I disagree heartily on both counts.

P. 103. The "correction for ties" in the case of the Wilcoxon paired sample case is not given.

P. 105. The Fisher-Yates normalized rank procedure is mentioned in passing, but without any description. There is confusion between power against normal alternatives and others.

P. 115 and following. This discussion of the Brown-Mood median tests for two-way analyses is based wholly on Mood's textbook and fails to mention an important subsequent article by Brown and Mood [1].

P. 123. The reference to Kendall's monograph on rank correlation is only to its first edition in 1948, without mention of the more recent and expanded edition in 1955 [3].

Comparative Contents. Tate and Clelland discuss two general areas that are not discussed by Siegel. First, nonparametric confidence intervals, and, second, short-cut normal theory methods. Two specific topics in the Tate and Clelland book and not in the Siegel book are (1) the Tukey-Olmsted corner test of association, and (2) the Brown-Mood analogs, based on counts above and below medians, to linear hypothesis problems. Siegel considers the Moses extreme reaction test, while Tate and Clelland do not. Siegel discusses permutation tests on the observations in more detail than do Tate and Clelland. Neither book discusses runs of kinds other than above and below the median and the Wald-Wolfowitz two-sample runs test. Neither book discusses in any detail problems and approximations connected with the Poisson distribution. Both books are weak in their discussions of power, and both omit discussions of design problems.

Tables. The tables and graphs given by Tate and Clelland are much the same as those given by Siegel, except that Tate and Clelland in general give critical values for a larger number of significance levels, typically 1%, 5%, 10%, and 20%. This is a distinct advantage, since one usually wants to know at least an approximate P-value, not just an accept-reject dichotomous decision. Also, Tate and Clelland include the Link-Wallace tables for short-cut testing and multiple comparisons in the balanced one- and two-way analysis of variance cases, along with Mosteller's graph for short-cut estimation of a normal correlation coefficient.

A first glance at many of the tables shows scattered blank spaces that may puzzle some readers. For example, in Table E, two-sided critical values are given for testing a null hypothesis on a binomial proportion in terms of number of observed "successes." For sample size 25 and null hypothesis proportion .5, the table reads as follows:

Significance level	*Critical values*
.20	8, 17
.10	
.05	7, 18
.01	5, 20

Why the blank for significance level .10? The answer is, of course, that because of discreteness, if one wants a significance level $\leq .10$, one actually obtains a level $\leq .05$ and this might as well be recognized. However, the point should be, and is not, made clear for the intended reader of the book.

Table D (confidence intervals for 50%, 75%, and 90% quantiles) and Table E (explained above) overlap in large measure, but they are not everywhere consistent. For example when sample size is 50 and significance level .2, we may compare the first column of D with the last column of E. Both cannot be correct. (In the last column of E, with $N=40$, $\alpha=.05$, and $\phi=.5$, "17" is a misprint for "27.")

In fact four of the tables (D, E, N, and O) all give in effect critical values for the sign test, but this extensive overlap seems to be mentioned only once (p. 35) very briefly. A superficial check turns up discrepancies beyond the one mentioned above. For example, with sample size 30 and two-tail significance level .2, Table E tells us that 11 is the critical number of successes in the left tail, while Table O shows a blank implying that 10 is critical. (The actual two-tail P-values are .2005 and .0987 respectively.) Again, for sample size 50 and two-tail significance level .2, Table O tells us that 20 is the critical number of successes at the left, while Table N gives 19 as the corresponding critical value. (This is for the same situation as in the above paragraph; Table E gives 20 and Table D gives 19. The actual two-tail P-values are .2026 and .1189 respectively.)

This is an appropriate place to interject a plea that tables of critical values for discrete statistics include P-values. The extra room needed is often a small price to pay for the added information.

Tate and Clelland give a compact table of approximate critical values for the 2×2 contingency table test covering a wide range of sample sizes. Siegel gives a much more detailed table for small sample sizes and no table for larger ones.

Tate and Clelland give two-tail critical values for the Wilcoxon signed rank test with sample sizes 4(1)30 and significance levels .01, .05, .10, and .20. Some of the values for level .20 appear in fact to correspond to P-values slightly greater than .20.

Table Q, giving critical values for the H test, has many errors, mostly small, but a few of consequence. Some of these stem from taking, say, a value of H as critical at the .01 level when the actual P-value is .011. Others appear to come from lack of reference to the correction note published by Wallis and me as a supplement to our paper presenting the H test [4]. One or two appear to be typographical errors. Siegel's table for the H test is much more complete and I have noted no errors in it.

Tate and Clelland give clearly drawn charts of the Clopper-Pearson kind for tests and confidence intervals on binomial proportions. Significance levels are .01, .05, .10, and .20 (two sided).

REFERENCES

[1] Brown, G. W., and Mood, A. M., "On median tests for linear hypotheses." Pp. 159–66 in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. Jerzy Neyman. Berkeley: University of California Press, 1951.

[2] Daniels, H. E., "A distribution-free test for regression parameters," *Annals of Mathematical Statistics*, 25 (1954), 499–513.

[3] Kendall, Maurice G., *Rank Correlation Methods*, Second Edition. London and New York: Charles Griffin and Hafner, 1955.

[4] Kruskal, William H., and Wallis, W. Allen, Errata Section, *Journal of the American Statistical Association*, 48 (1953), 907–11.

[5] Moses, Lincoln, Chapter 18, "Non-parametric Methods," pp. 426–50 in *Statistical Inference* by Helen M. Walker and Joseph Lev. New York: Henry Holt, 1953.

[6] Patnaik, P. B., "The non-central χ^2- and F-distributions and their applications," *Biometrika*, 36 (1949), 202–32.

[7] Savage, I. Richard, "Nonparametric Statistics," *Journal of the American Statistical Association*, 52 (1957), 331–44.

[8] Siegel, Sidney, *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.

The Mathematical Theory of Sampling. *Walter A. Hendricks.* New Brunswick, N. J.: The Scarecrow Press, 1956. $7.50.

T. A. BANCROFT, *Iowa State College*

THE author states that this book is addressed to students who wish to become familiar with formulas and computations incident to sample design and the analysis of sample data. He states further that although major emphasis is given to sample surveys, most of the methods have a broader field of application. The book takes up applications largely from the viewpoint of agricultural surveys "because that happens to be the author's field of activity."

Chap. I presents an introduction to basic concepts in sampling and discusses the role of statistical methods in sampling. Chap. II deals with the case of random sampling with equal probabilities of selection for infinite and finite universes. Expected values, averages, and standard errors are presented. Chap. III takes up probability

theory and its relation to sampling including a discussion of Boyes' Theorem, binomial variance and approximate confidence limits, sampling without replacement from a finite universe, and variances and covariances in the binomial and multinomial distributions.

Chap. IV covers random sampling in a universe defined by a continuous frequency distribution with emphasis on derived sampling distributions based on a normal parent population distribution.

Stratified sampling is the subject of Chap V which includes a discussion of optimum sample allocation, the relationshiop of costs to sample allocation, and the effects of errors in stratum weights.

The author takes time out in Chap. VI to present a rather detailed discussion of the analysis of variance including: equal and unequal numbers per stratum, data grouped in subclasses, two-dimensional stratification of data, two-way classifications with unequal cell frequencies, and studies of attributes.

The remaining chapters treat the usual topics in survey sampling: Chap. VII Multistage Sampling, Chap. VIII Sampling Units and Their Selection, Chap. IX Sampling with Unequal Probabilities of Selection, Chap. X Ratio and Regression Estimates, and Chap. XI Control of Errors in Sample Surveys.

Numerous numerical and theoretical examples are given which should prove helpful in understanding the derivations and exposition; however, no exercises are included. References given at the end of the chapters provide a fairly satisfactory bibliography in survey sampling, and also in general statistical methods and theory up to the intermediate level. The one exception noted is the omission of any reference to *Sampling Theory of Surveys with Applications* by P. V. Sukhatme.

The author has combined a theoretical approach involving some mathematical literacy with numerous simple numerical examples to illustrate the ideas and to indicate the nature of applications. Also considerable space is devoted to topics in general statistical methods and theory which other texts in survey sampling assume as prerequisites. As a consequence the book is hardly suitable as a beginning methods course in survey sampling of the type of Volume I of *Sample Survey Methods and Theory* by M. H. Hansen, W. N. Hurwitz, and W. G. Madow. On the other hand, no attempt is made to present a complete formal mathematical treatment and hence, in this sense, it cannot be classed as an advanced text. Nevertheless, this reviewer recommends this book to students of survey sampling who have some knowledge of mathematics through calculus and who wish a self-contained treatment of the subject including related topics in general statistics at the intermediate level.

Extension of Existing Tables of Attributes Sampling Plans. *J. M. Wiesen.* Albuquerque, New Mexico: Sandia Corporation, 1956. Available as "Physics (TID-4500, 13th Edition)," Office of Technical Services, Department of Commerce, Washington 25, D. C. Pp. iii, 236. $1.25. Multilith. Paper.

DAVID R. HOWES, *Chemical Corps Engineering Command*

THIS document presents 230 curves for each of the acceptance numbers 0, 1, and 2, for sampling plans with lot sizes of 100 to 500 by steps of 50 and from 500 to 2000 by steps of 100. Sample sizes are up to one half of the lot by an irregular rule starting at 25 for a lot of 100, and 200 for a lot of 2000. Since the sample sizes are large with respect to the lot size, hypergeometric probabilities are used. The charts can be read to an accuracy of .05 per cent in per cent defective and to .2 per cent in probability of acceptance.

The report is designed to provide operating characteristics for single sampling plans of very high discrimination for use where sequential or variables sampling is impractical. It may also be useful as a source of hypergeometric probabilities as an

alternative to the use of the approximations given by Dodge and Romig as formula (1′), p. 43, in *Sampling Inspection Tables* (Wiley, 1944). For practical purposes, it would be helpful to have, for given sample sizes, charts showing selected percentage points of the hypergeometric distribution as a function of lot size.

Tables of the Normal Probability Integral. D. B. Owen. Albuquerque, New Mexico: Sandia Corporation, 1957. Available from the Office of Technical Services, Department of Commerce, Washington 25, D. C. Pp. 58, $0.40. Paper.

HANS O. BREMER, *National Bureau of Standards*

THIS publication was prepared primarily as a working paper for use within the Sandia Corporation. It contains tables of the normal probability integral and Mill's ratio.

The form of the normal probability integral that was tabulated is

$$G(h) = (1/\sqrt{2\pi}) \int_{-\infty}^{h} e^{-1/2x^2} dx$$

and is tabulated for $\pm h = 0(0.001)4(0.01)7$ with the function being given to eight decimal places. For extreme values of h, the normal probability integral $G(-h)$ and Mill's ratio are given for $+h = 50(1)150(5)500$ to eight significant places.

These represent an expansion of previous tables given by E. S. Pearson and H. O. Hartley in *Biometrika Tables for Statisticians*, Vol. I, Cambridge University Press, London, 1954, pp. 104–11.

Metzger's Table of 10,000 Square Roots. *William T. Metzger.* Indianapolis, Indiana: Research Tabulating Bureau, 1958. Pp. 21. $2.50. Paper.

THIS table shows each number from 1 to 10,000 and its square. The square root of a given number is obtained by finding the location of the given number in a column of squares, and then noting what two four-digit numbers produce squares on either side of the given number.

The table is reproduced from tabulating machine print-outs in rather small type. It consists of 100 columns of 100 entries each.

W. A. W.

Metzger's Table of 45,000 Per Cents. *William T. Metzger.* Indianapolis 6, Indiana: Research Tabulating Bureau, P. O. Box 7, Dept. 434, 1957. $4.00. Paper.

DENOMINATORS from 1 to 300 are represented by columns and numerators from 1 to the value of the denominator are represented by rows. The entries then give the percentages to five decimals.

The table is reproduced from tabulating machine print-outs in rather small type.

W. A. W.

Models of Man. *Herbert A. Simon.* New York: John Wiley and Sons, 1957. $5.00.

MAY BRODBECK, *University of Minnesota*

THE descriptive phrase "Mathematical essays on rational human behavior" is helpfully appended to the title of this book. More specifically, these papers on diverse topics in social science, exclusive of economics, are presented as a contribution to quantified theory of the social psychology of groups and organizations. Their unity, as the author conceives it, lies in how much they contribute toward "laying foundations" for a science of man as a social and rational animal. Whether or not

they actually have this unity is doubtless just another instance or facet of the aggregation problem. My own reaction to the aggregation is ambivalent. Since Simon's motives are admirable and his ingenuity is considerable, this qualified enthusiasm requires some explanation.

The essays are divided into four sections. Each section has a preface in which Simon discusses the connections among the parts and elucidates some of the special problems raised. Many of these remarks are, as one would expect, wise and illuminating comments on problems of measurement and quantification in social science. Here indeed is one source of the reviewer's ambivalence. One would as soon censure Motherhood as the ideals Simon professes. Scientific terms must be operationally defined. Of course. Quantification is highly desirable—and not just because physicists do it. Of course. A truly unified science of human behavior is devoutly to be wished for. Indeed. Nor, given the present state of social science, outside of economics and certain branches of theoretical psychology, is it superfluous to rehearse these maxims. Indeed not. Yet, two things trouble me. The first is Simon's misplaced preoccupation with the analysis of causality. It is misplaced because, among other things, he quite incorrectly, it seems to this reviewer, believes such an analysis to be necessary for a proper operational definition of "power" or "influence" in the social or political sense. The second qualification is the "level" of theory at which Simon constructs his major "unifying" model. I shall discuss these one at a time.

The first section is on "Causation and influence processes." The very concatenation of these two notions is odd. Causality is a methodological notion, while influence between persons, as in "Smith influences Jones," is a natural or descriptive relation like "loves" or "taller than." In the first paper, "Causal ordering and identifiability," Simon develops a method for establishing a causal order among the variables of a theory and a corresponding order of "precedence" among the equations containing these variables. Causal imputation is given, according to Simon, "operational" meaning by associating with each equation a procedure for altering its constant terms or coefficients. Econometricians will doubtless be interested in the parallelism Simon draws between the conditions for establishing a causal order and those for the identifiability of parameters. However, "operational definition" is a good thing, all right, but only where appropriate, namely, to descriptive or empirical concepts or variables, particularly those involving measurement. "Causality" is not a descriptive concept, that is, a name for a kind of observable thing, event, or attribute of or relation between things or events. The problem of analyzing "causality" is thus not a problem of proper concept formation nor, *ipso facto*, of "operational definition." It is, rather, part of the problem of knowing when we have a true empirical law or correlation and when we have an accidental conjunction or a "spurious" correlation. On *this* problem, Simon has useful things to say in a separate essay, about the assumptions scientists make when distinguishing true from spurious correlations. These assumptions imply the need for a context of knowledge, that is, a theory about the irrelevance of certain variables or the lack of correlation among random disturbances associated with given relevant variables. They do not imply or require an "operational" definition of causality, for there is no such thing. There is no such thing because "causality" is not a variable of the law or theory we are testing. Simon knows all this, of course, so he speaks instead of the "operational definition of an equation." By this he means associating with the equation a procedure for manipulating the coefficients. Unfortunately, I do not know what it means to define an equation, operationally or otherwise. For one thing, a definition is unequivocal, while the ways variables can be manipulated may be legion. I do know that this over-extension of "operational definition" to places where it is not applicable is more confusing than clarifying.

One consequence of this confusion is Simon's belief that the definition of influence requires a prior definition of causality. In the same section is reprinted a paper on the "Observation and measurement of political power." Since "power" or "influence," which he uses synonymously, is an asymmetrical relation between persons, Simon fears that, like "cause," it may be a suspect notion. But, unlike "cause," "influence" is a descriptive relation, hence it can be measured, correlated with other things, and so on. It is really not necessary to point, as he does, to such recondite things as electrical network theory in order to justify the use of asymmetric relations within science. Is "harder than" something to be "wary" of? The characterization of power or influence no more depends upon a prior analysis of "cause" than does the definition of "harder than." An operational definition of "Jones influences Smith" or "Jones' behavior causes Smith's behavior" might be "If Jones talks, then Smith jumps," just as "harder than" is defined in terms of scratching. No theory of causality nor even of the difference between a true and a spurious correlation is needed for an operational definition. Needless to say, Simon's practice is better than his theory, as is shown by his subtle discussion of the problems in constructing a useful definition of "influence" and in the choice of units of observation and measurement.

Section II, "Social processes," deals with man as a social animal. These models are "mathematical" or, as I prefer to say, quantitative reformulations of nonquantitative theories of the behavior of individuals in groups—their communication processes, the stability of the group, etc.—as developed by the social psychologists Homans and Festinger. Such "translations" are not only, as Simon points out, clearer and more rigorous, they also say a great deal more than the original. This is both a virtue and a weakness. The virtue is obvious. The weakness is that by the nature of the mathematical machinery and the data, the quantified theory will generally, even if not in these particular cases, say more about fewer variables. This is partly because of the number of equations that can be handled at one time and partly because many of the relevant variables resist quantification. A nonquantified theory, on the other hand, may say more, though less precisely, by stating the connections among a greater number of variables. One may therefore purchase the virtues of quantification at the price of theoretical triviality. This need not happen, of course, as theoretical economics shows. But I shall revert to this case in a moment. In contrast with the models of "social" man as an adaptive being who engages in group activity for its own sake, the two papers in Section III, headed "Motivation: inducements and contributions," present models of man as an economic maximimizing animal, for whom organizations are purely instrumental to external rewards. The concepts of classical economic theory are here applied to theories of organization and of the employment relation. The final, fourth section, on "Rationality and administrative decision making" is in many ways the most interesting part of the book. It attempts to unify in a new theory certain features of social and rational man. Simon here argues that a "more realistic" definition of rational behavior is needed, one which takes account of the empirical limits, within the individual and in his environment, on human rationality. Economic, maximizing, omniscient man gives way to "satisficing" man of limited knowledge who settles for a course of action that is "good enough" rather than the best of all possible alternatives—which in any case he never knows. Simon's definition of "approximate" or "intended" rationality incorporates social and psychological variables omitted from both classical economics and more recent game and statistical-decision theory. These variables, drawn from "common experience," help provide a definition of rational choice which is more consonant with how people "actually" do behave.

Quite apart from any difficulties in recent economic theory, Simon's attempt to replace the unrealistic "ideal type" of economic man by something more true to life

seems inherently desirable, on the face of it. Yet, I am less persuaded of the virtues of this project than perhaps a good empiricist should be. Thus far, the most fruitful theory-construction in social science has been either of two kinds. At one extreme, as in learning theory in psychology, the behaviors studied are of the rudimentary kind exhibited by, say, the white rat. (This is why Simon must rely on conjecture for the variables in "actual" human rational behavior.) Quantification is possible, but prediction of complex, human behavior waits upon the development of composition laws, that is, laws stating how the laws of the elementary processes combine to result in complex human behavior, e.g., so-called "insight" learning in man may be explained in terms of composition laws based upon the laws of rudimentary trial-and-error learning. If, as there is reason to believe, human behavior consists of complex patterns of elementary processes, the potential scope of such theories is very great. At the other extreme, in theoretical economics, highly complex and sophisticated aggregate human behavior is explained by means of assumptions about rational behavior which, for any individual, are most likely false. Yet, for macrophenomena, they work rather well. Unlike either learning or economic theorists, Simon attempts to construct directly a theory for individual choice behavior. This, at the level of social psychology, is neither very basic nor very macroscopic. Simon has the advantage over most social psychologists of mathematical ingenuity and understanding of what has already been done in economics. Whether the probable lack in either actual or potential scope of theory which is neither fundamental nor highly aggregative can be compensated for by these advantages remains to be seen.

The Demand and Supply of Scientific Personnel. *David M. Blank and George J. Stigler.* New York: National Bureau of Economic Research, 1957. Pp. xix, 200. $4.00.

M. Bronfenbrenner, *Michigan State University*

This is the most recent and should prove one of the most controversial of the National Bureau studies in service industry labor markets. The series is a monument to one of the present authors, George J. Stigler, who has been associated prominently with every volume.

Unfriendly critics have labelled the National Bureau a "figure factory" in the tradition of "Knowledge for what?" Such charges are definitely out of place with respect to these service industry labor market studies. Chapter and verse defy detailed citation, but the studies all stress the beneficent operation of the free market in the long run, adapting the quantity supplied to changes in demand, or vice versa, via the price mechanism.

The same invisible hand is equally visible in the present unusually timely study. Under the general head of "scientific personnel" major emphasis is placed on the market for the services of engineers, who dominate the category quantitatively and for whom the most and best data are available. Secondary stress is on chemists. Mathematicians and physicists are dealt with together in Chap. 5, and there are a number of useful but scattered references to the problems of "earth scientists" (geologists and geographers). Mathematical statisticians are not considered separately from other mathematicians, and other statisticians are excluded entirely. Despite the sharp increase in employment opportunities since 1939 in all the fields covered, supply has kept pace, with possible dilution of quality, at prices rising proportionately only half as much as the wages of manufacturing labor. When shortages were most acute (1946 and 1950) the incomes of scientific personnel were lagging furthest behind, and reduction of the lag alleviated the shortages.

It may be informative to reproduce some of the authors' indices, in view of the widespread belief that incomes of scientists have risen inordinately. Table 11 (p. 25)

compares indices of median engineering salaries to average wages, salaries, or net incomes in a number of other occupations. Taking the 1929 value of each ratio of median engineering salary to the corresponding figure for another occupation as 100, the most recent values computed by Blank and Stigler are: to all full-time wage and salaried employees, 67.9 (1954), to full-time manufacturing wage earners, 66.6 (1954), to physicians, 59.5 (1951), and to dentists, 83.5 (1951). Only the ratios to lawyers (96.0—1954) and to college teachers (96.3—1953) show the engineers holding their own. If 1939 instead of 1929 had been chosen as a base year, all the comparisons would have been less favorable to the engineers, although in varying degrees.

A statistically interesting feature of Blank and Stigler's analysis explores the tendency for scientific personnel to appear when needed "out of the woodwork," meaning that increases in their number are more rapid than can be accounted for by the net output of accredited training institutions. Anyone who has kept his ears open around, say, an aircraft or electronics plant knows what has happened, namely upgrading, particularly the application of the "engineer" title to intelligent technicians trained on the job. But this reviewer at least did not realize the number of such cases—apparently some 40 per cent of the engineering profession (p. 86).

The implicit advice to trust the market, keep one's price system dry, and resolve all shortages by income increases, is not currently popular as regards scientific personnel, thanks to recent Soviet advances (ICBM, Sputnik). One cannot, however, criticize this volume's pre-Sputnik complacency as wrong; at the very worst, it may be partially irrelevant. It shows how the market equates the supply of scientists to the demand. The arguments of the alarmists are that the supply and the demand are both too low to keep pace with the Soviet Union, and that the price system allocates the supply wrongly (too many engineers in gadgetry and salesmanship, too few in research and teaching). Blank and Stigler's careful statistics provide no clues for evaluation of these assertions. But neither does any alternative body of scientific analysis to supplement sheer intuition.

New Technology and the Supply of Petroleum: the Treatment of Uncertainty in Resource Planning. *Miller B. Spangler.* Chicago: University of Chicago Press, 1957. Pp. xiii, 299. $3.50. (Obtainable from the author, R.F.D. 1, Yorktown Heights, New York.)

JOHN S. MCGEE, *The University of Chicago*

MUCH has been said about the uncertainties that enshroud crude oil discovery and production. The first successful oil well in the United States was drilled in 1859, and periodically ever since controversy has developed over the volume and adequacy of domestic oil reserves. Many predictions have been made. Those that could be tested have generally proved to be grossly inaccurate.

The purpose of this volume is to develop better methods for predicting the discovery and development of aggregate crude oil reserves in the United States. Orthodox methods, including projection of time series, have not been notably successful. What is wanted, according to the author, is a new technique for dealing with complex uncertainties. The solution is to rely upon expert opinion. Experts can prepare subjective probability estimates for each expected technological innovation, together with estimates of its impact upon domestic oil reserves. Nontechnological factors are taken to be constant.

Underlying this approach is, among other things, a special conception of innovation itself. These are the basic tenets: (1) Innovation is accelerating. (2) Specific innovations are more or less predictable, since there are "orderly steps proceeding from fundamental research . . . to developmental research" (p. 7). (3) Because laws of

technological development are analogous to biological laws, experts can determine the relative maturity of specific techniques and processes, and how much additional development is possible in various fields. (4) It is possible to estimate what kind, as well as how much, innovating there remains to be done, and what effects it will have.[1]

By many standards, the United States is a relatively mature oil province, and the physical task imposed upon discovery and development has grown harder. Nevertheless, technology has largely filled the gap. As a consequence, the real price of domestic crude oil has remained "rather stable for the past several decades." (p. 44). The outlook is for continuing technological progress, since the industry is increasing what are already large research expenditures.

The basic question posed by this study is, in a sense, whether there will be enough innovation. Lacking the probability estimates of a corps of experts, the author hints at his own. He distinguishes three main areas for future technological development: discovery, field development, and recovery. For each he discusses, in considerable detail, the most significant past innovations; appraises the relative maturity of existing techniques; and cites present "problem areas" which are peculiarly attractive to innovation. He then reaches some broad conclusions about future domestic oil reserves, insofar as they are technologically determined. He concludes that "the supply of domestic petroleum through 1975 appears to hold good promise of meeting demand [at approximately present prices] . . ." (p. 293).

This study raises many questions. The author stresses the need for a method of prediction that is both "practical" and superior to the old techniques. The operability, as well as the meaning, of the one he suggests is at least questionable. Neither is it clear that guesses based upon guesses will be more useful than those orthodox techniques that the author eschews. Furthermore, the author does not present a systematic method for synthesizing (and interpreting) disparate subjective estimates of different expected innovations. There are also some missing steps that make this treatment difficult to follow. For example, the analysis is directed primarily to estimating reserves at some future time. But reserves is a stock which has been large, but not entirely stable, relative to yearly additions and withdrawals. The analysis of price requires some notion about the withdrawals from stock per unit of time, and the relationship to reserves is not spelled out. Finally, there are some observed facts that call the main conclusions into question, though they do not necessarily overturn them. There is evidently a great deal more to world petroleum supply than technological progress, for domestic innovation has not so far proved a match for advantages enjoyed by giant foreign fields that now contain the bulk of the world's known oil reserves. In a free market, decisions made by many independent firms may be a better gauge of probabilities than any amount of interviewing of experts. In spite of restraints upon world petroleum trade, in recent years domestic oil has retained a declining share of the U. S. market. The recent modest record of net additions to reserves as compared with past performance, and the reaction of domestic companies to it, suggest that the industry now has reason to be less optimistic about domestic oil prospects than is the author.

Nevertheless, this is an interesting book. The author brings to it an uncommon combination of skills and interests. He knows a great deal about petroleum technology, and usually does a good job of communicating what he knows. Anyone who

[1] These notions, and related ones, are not wholly novel. Indeed, they have long enjoyed a certain popularity. Compare, for example, Report of the United States National Resources Committee, Technological Trends and National Policy, 1936; Hornell Hart, "Technology and the growth of political areas," Technology and International Relations, ed. by William Ogburn, Chicago: University of Chicago Press, 1949.

The predictability of invention has recently been questioned vigorously. John Jewkes, David Sawers, and Richard Stillerman, *The Sources of Invention;* New York: St. Martin's Press, 1958.

is interested in the petroleum industry can find something stimulating here. Although the data assembled are primarily from secondary sources, and the references favor the trade journals rather heavily, they form a good introduction to petroleum technology.

Distribution of Union Membership among the States, 1939 and 1953. *Leo Troy.* Occasional Paper No. 56. New York: National Bureau of Economic Research, 1957. Pp. viii, 32. $.75. Paper.

MARGARET S. GORDON, *University of California (Berkeley)*

THERE has long been a need for statistics of union membership by state or region. In the absence of reliable data of this kind, it is difficult to determine what geographical shifts in union strength are taking place or to analyze the factors responsible for these shifts. Yet the problems involved in the compilation of reliable data are formidable, since only two states collect statistics of union membership, and it is difficult to allocate membership accurately by states from the records that can be furnished by some two hundred national unions.

The present report represents a valuable contribution toward filling this gap. The author was able to obtain records of local union membership from almost all national unions, as well as the cooperation and assistance of the national unions in determining how to allocate the membership by states. Because the job was time-consuming and costly, the estimates were limited to the years 1939 and 1953, but the results provide a basis for analyzing geographical differences in rates of growth during a period of rapid expansion in union membership.

Despite their usefulness, the statistics must be interpreted with caution. The absence of data on local independent unions represents an important gap, since the membership of these unions is unevenly distributed by states. Some of the other difficulties in interpretating the data are, if anything, understated by the author—the problems of defining union membership, of duplication of membership, and of allocating the membership of local unions which spread over several states, to mention only a few.

In some respects, the difficulties presented by union membership data obtained from unions are analogous to the problems presented by employment data obtained from establishments. As Irving Bernstein has pointed out,[1] it would be extremely valuable to have union membership data obtained from the decennial census, or at least from a census-type survey. Not only would problems of double-counting and dubious allocation by state or region be eliminated, but it would also be possible to anlyze the data by age, occupation, and other significant variables on which information is not usually available from union records. Clearly, also, the growth of union membership could be analyzed much more effectively if other states would inaugurate data-collection programs similar to those in effect in California and Massachusetts.

Residential Finance, 1950. *Richard U. Ratcliff, Daniel B. Rathbun, and Junia Honnold.* For the Social Science Research Council in cooperation with the U. S. Department of Commerce, Bureau of the Census. New York: John Wiley and Sons, Inc.; London: Chapman and Hall, Ltd., 1957, Pp. x, 180. $6.00.

SAUL B. KLAMAN, *National Association of Mutual Savings Banks*

INVESTIGATORS who have had occasion to use the mass of detailed statistics on residential financing in Volume IV of the 1950 Census of Housing will welcome the summary and analysis of these statistics presented in *Residential Finance, 1950.*

[1] Irving Bernstein, "Measuring with a broken yardstick," in *Statistics of Labor-Management Relations*, Institute of Industrial Relations, University of California, Berkeley, 1956, pp. 18–19.

This monograph is the seventh in a series of thirteen which have been, or soon will be, published on the 1950 Census of Population and Housing.

The authors have done an admirable job of reducing over 600 pages of elaborate Census tables to 176 pages of easily readable text and summary tables. They have organized the vast amount of Census material into six well planned chapters (plus a brief introductory chapter) on over-all mortgage characteristics, lender practices, borrower characteristics, junior mortgage financing, financing of rental housing, and government mortgage guarantees. This arrangment and choice of subject matter is well suited to bringing out the main factual findings of the 1950 Census survey of residential financing.

The reader should expect little more than a factual summary, however, for as the authors themselves warn "we have deliberately chosen to favor extensive treatment at the expense of intensive refined analysis (and hence cannot avoid) the tiresome recitation of fact inherent in extensive treatment. Detailed examination of a limited number of arbitrarily selected hypotheses would lend analytical elegance to our work . . . (but) would run the risk of ignoring more important material." This reviewer feels that there is a middle position which would have permitted some refined analysis in significant areas of mortgage finance within an over-all framework of extensive treatment. The "risk of ignoring more important material" is ever present in the work of the economic analyst and should not deter him from selecting, on the basis of knowledge and experience, those basic hypotheses which he considers worthy of intensive analysis.

One of the shortcomings of adhering too closely to a recitation of Census facts is the absence in this monograph of any comparative analysis of pertinent data from other sources for years since 1950. Such analysis would have been particularly relevant in connection with some of the material in Chap. 3 on lender practices, not only because current data are readily available in this area but also because there have been such broad changes in mortgage portfolio composition since 1950 that the discussion of Census data is of historical interest only. Indeed, analysis of Census findings as if they had some currency, in a monograph published seven years after the date of reference, may tend to mislead the reader.

In their discussion of mutual savings banks, for example, the authors state that "participation in both the FHA and VA programs is less extensive than in the case of commercial banks. . . . In terms of dollar volume, their holdings of conventional first and junior mortgages amount to two-thirds of outstanding mortgage debt" (page 26). Basic legal and economic changes since 1950 have markedly altered savings bank mortgage practices so that they have placed increasing emphasis on federally underwritten mortgages and in recent years have had a substantially larger proportion of their portfolios in these loans than commercial banks have had. Conversely, the proportion of savings bank holdings in conventional mortgages has been sharply reduced, and is now less than that of commercial banks.

Similarly, the discussion of borrower characteristics in Chap. 4 would have been of considerably more interest and use if the 1950 Census results were compared with more recent findings of the sort included in the Federal Reserve annual Survey of Consumer Finances. Data for income and occupation of mortgage borrowers and for market value and outstanding indebtedness on owner-occupied homes, similar to those reported in the 1950 Census of Housing, have been available from the Survey of Consumer Finances for several years. Lack of complete comparability between these data might serve to qualify but not to invalidate conclusions.

For areas in which later data are not available, it would have been useful for the authors to have commented on important market changes which have taken place since 1950 and which, in their judgment, have rendered obsolete for current analysis

some of the Census findings. Token recognition of basic changes was given in connection with the discussion of junior mortgage financing: "Since the time of the enumeration of the 1950 census, there have been changes in the mortgage market which, no doubt, have altered the incidence of junior financing" (p. 70). Comments of this type, reinforced by brief analysis, would have been pertinent to several other sections of the monograph.

A point on which I cannot resist comment, though it is minor in the framework of the monograph, is that the authors have failed, unfortunately, to make a conceptual distinction between techniques of mortgage acquisition and secondary market transactions. They have, in effect, accepted the notion, common among mortgage market participants, that the acquisition of any mortgage by an investor from an originator is a secondary market transaction even though the mortgage may have been newly originated on the basis of a firm commitment from the investor to purchase it. It is this special meaning which the authors have in mind when they cite the large volume of mortgage "purchases" by life insurance companies as evidence of their heavy dependence upon the secondary market. The bulk of such "purchases," however, are not secondary transactions under the accepted economic concept of a secondary market—one in which trading takes place between investors in existing mortgages or securities.

Most so-called "purchases" by life insurance companies merely reflect the indirect origination of loans through agents or mortgage correspondents, a lending technique common to investors engaged in a nationwide mortgage program. In the economic sense, such acquisitions are primary market transactions just as are acquisitions of mortgages by life companies through their own branch offices, or direct origination of mortgages by lenders in local areas. The authors would have improved understanding of mortgage market techniques and processes if they had discussed the Census material simply in terms of methods of acquisition rather than in terms of secondary and primary market activity. It is not possible from Census data, unfortunately, to determine the volume of "true" secondary market activity. It is generally known, however, that life insurance companies and mutual savings banks are more actively engaged in the acquisition of seasoned mortgages in secondary markets than are commercial banks or savings and loan associations.

It would have been useful if the authors had discussed the reliability of Census data in cases where results clearly do not agree with common understanding of institutional arrangements or market operations. The Census finding that mortgage companies are active purchasers of federally underwritten loans, for example, does not square with other information. It is not likely, either, in the light of recent findings, that they held as large a proportion of conventional mortgages in 1950, as indicated by Census data. (See my article on "Mortgage companies in the postwar mortgage market," *Journal of Finance*, May, 1957).

The relatively large percentage of high loan-to-value ratio conventional mortgages held by commercial banks and other institutional lenders, in some cases higher than permitted by statute, is admittedly unexplainable by the authors. Even allowing for liberal property appraisals, the reported finding is suspect that 22 per cent of all commercial bank loans and 15 per cent of all life insurance company loans, made to nonwhite borrowers were in excess of 100 per cent of purchase price. In the absence of discussion on statistical reliability, other Census findings not so clearly out of line are brought under some suspicion.

The shortcomings of the monograph are not so many or so important as to impair its general usefulness. The authors have achieved their stated purpose, that of presenting an extensive summary of the mass of tabular material in the 1950 Census

volume on residential financing. They have covered a number of basic topics in a highly competent fashion and in a useful organizational scheme. As a reference work, this monograph will be a valuable tool for investigators.

Consumer Instalment Credit, Part II, Vols. 1 and 2. *Conference on Regulation. Board of Governors of the Federal Reserve System.* Washington, D. C.: U. S. Government Printing Office, 1957. Pp. xxvi, 553; x, 161. $1.75; 0.60. Paper.

VIOLA WYCKOFF, *New York University*

THESE volumes are concerned with the case for and against selective control of consumer credit. The second volume takes up the pros and cons. The first volume considers consumer credit within our economy: its relation to economic stability and instability, wartime regulation, quantity and quality of consumer credit during the cycle, effect of general credit controls on consumer credit, borrowers, marketing practices, etc.

The proceedings of the conference are composed of papers and comments by 34 economists including Thomas Atkinson, Milton Friedman, Don D. Humphrey, F. A. Lutz, Geoffrey Moore, Robert Shay, and Arthur Smithies. The discussions are provocative, and the first volume is comprehensive.

The important questions considered in the second volume were: Does general credit control affect consumer credit? Are general controls adequate to control consumer credit? Is supplementary, selective control of consumer credit needed? Are adequate laws possible and feasible now? Would consumer credit regulation discriminate against individuals and enterprises?

These questions were not settled. Concerning the first two question Eli Shapiro and David Meiselman stated, "The relatively fixed proportion of total bank loans that consumers and finance companies accounted for appears to contradict the assertion, at least for the 1953–55 period, that banks so favor providing funds to finance consumers that traditional weapons of monetary policy can have little effect on bank sources of consumer finance. On the contrary, banks respond to central-bank policy by altering the availability of funds both to consumers and to other borrowers, among them business borrowers" (1, 308–9).

Marcus Nadler disagreed: "From 1946 to 1955, according to a recent Federal Reserve survey, member bank loans to sales finance companies increased twice as rapidly as all business loans. In addition, these companies were charged a lower rate of interest than any other business group. Even in periods of tight money, they [sales finance companies] have little difficulty in obtaining both short- and long-term funds, as was strikingly demonstrated in 1955" (2, 3).

Practically everyone viewed the question of consumer credit regulation from a different angle, even when their conclusions were the same. Some economists agreed about arguments and facts but drew different conclusions, for example, Nadler and Lester Chandler. The latter opposed control and thought that " . . . attempts to impose controls on consumer credit might well weaken the prestige and power of the central bank to such an extent that it could not use effectively even its general restrictive powers" (2, 30).

There was little disagreement on the importance of consumer credit within the economy; its relation to economic stability or to inflation. Those who favored control based their arguments mostly on these considerations. Those who opposed it preferred the operation of a "free market," and thought that controls would be discriminatory. (Some economists were fascinated by the folklore on thrift and tried to disentangle their opinions as economists from their bias as savers or borrowers. Ethical considerations cropped up throughout the papers.)

The type of law which might best control consumer credit and how to administer it were referred to only in passing. The other shortcoming to this reviewer arose from the topic selected, that is from consideration of consumer credit separately from other types of credit, not as a part of all credit within the economy. Federal regulation, if concerned with quantity of consumer credit in relation to the economy, would need to dovetail consumer credit regulation with other credit regulation in order to be effective. State laws such as New York's recent ones might well take care of quality of consumer credit and marketing practices.

The arguments and material presented may not answer definitively the question of whether or not consumer credit should be regulated, but they do present powerful reasons why any agency might be reluctant to undertake the task.

Consumption and Business Fluctuations. A Case Study of the Shoe, Leather, Hide Sequence. *Ruth P. Mack.* New York: National Bureau of Economic Research, 1956. Pp. xvii, 293. $7.50.

DANIEL B. SUITS, *University of Michigan*

IT has always seemed to me that economic aggregation is largely a crude simplifying expedient to enable the economist to press forward his research in the face of data and method inadequate to handle the detail of underlying economic reality. It is this detail that is the stuff on which theory is properly built and the interrelations of which are of real scientific concern; aggregates are properly defined simply as the sum of the parts. Yet in our professional society prestige frequently seems to accrue in direct proportion to the extent to which detail is washed over by global totals, rather than the extent to which careful observation of many facts is welded into a unified structure. Thus a study of "Consumption and Business Fluctuations"—about as aggregative and global as you can get—is prestigeous, while the attempt to observe carefully, analyze and systematize the behavior of the "Shoe, Leather, Hide Sequence" must derive such reflected glory as it can by proudly attaching itself to the symbol of research prowess as a "Case Study."

I believe these values are backwards and wrong, and I think Ruth Mack believes so too, but I can see no other reason for the title of her book. It is not, I am happy to say, a study of "Consumption" or "Business Cycles" at all—at least not in the aggregative sense in which these terms are understood. Rather, it is a thorough and patient study of the time pattern of the stocks, flows, and values of the shoe, leather, and hide industry. It is a detailed study centered on the idea that "we need to examine a wide variety of data not limited to the flow of goods and prices. We need to improve our ability to utilize these records for the purpose of observing and describing behavior. We need to study objectives, techniques, evaluations and actions of business men, for only through such knowledge can useful hypotheses be formulated concerning the dynamics of interconnection among economic events" (p. 32).

The organization of the study proceeds from a summary and statement of objectives to a brief description of the industry and its organization. This is followed by a cyclical analysis—more or less along lines familiar as standard National Bureau practice—of the time series of retail and wholesale shoe sales, leather and shoe production, and hide movement into sight, as well as retail and factory shoe prices, and prices of leather and hides. It is shown that while wholesale sales tend to lead, the remaining physical flows move together. The problem for study is the exploration of these facts—to explain them by careful reference to the behavior of consumers and producers at the several stages of production.

The production and sales of the shoe, leather, and hide industry follow, of course, the general pattern of business fluctuations in the economy but particular interest is centered on the origin and propagation of industrial subcycles averaging about a

year and a half in duration. There were estimated to be over a dozen such cycles in the period 1922–1939 which the study covers. Comparison of the amplitudes and turns of these cycles at the several stages of production is integrated with study of business practices to give insight into the propagation mechanism.

The path of the investigation leads directly upstream: chapters 5 and 6 are devoted to consumer purchases, 7, 8, 9 and 10 to the role of the retailer, 11 and 12 to the leather buying behavior of the shoe manufacturer, 13, 14 and 15 to hide production, prices and tanning, and 16 and 17 to summaries of the findings.

At each stage the history of the variables is referred to the relevant facts of business conduct. More often this is done qualitatively, but sometimes with the aid of regression analysis, and always with complete responsibility and full recognition of the limitations arising from method and data.

As a contribution the study is a beginning rather than an end. It outlines the problems to be analyzed and points directions for next steps. This is not an easy book to read. Indeed the amount of positive effort which must be invested is well above average. Since the author is not a mathematical economist, both reader and writer must handle the analysis "in bulk" unaided by mathematical notation or equipment, yet the material dealt with surpasses in complexity and detail many more global models usually considered beyond the range of literary analysis. On the other hand the reader willing to invest the time and effort will be adequately repaid, for he will find himself treated, not as audience, but as peer and partner whose help is sought and opinions valued.

The Measurement of Levels of Living with Special Reference to Jamaica. Colonial Research Studies No. 24. *C. A. Moser, Reader in Social Statistics, University of London.* London: Colonial Office. Her Majesty's Stationery Office, 1957. Pp. 106. 13s. 6d. Paper.

ALICE HANSON JONES, *Washington, D.C.*

THIS concise booklet is much more than a presentation of selected data. It is a dexterous pilot study, sponsored by the Colonial Economic Research Committee and based on a nine weeks stay in Jamacia in the summer of 1955, of what is involved in a specific attempt to implement recommendations developed within the United Nations organizations.[1] The statistical methods are not involved, but the conceptual problems are intricate. Though focused on one under-developed colonial area at the present time, and though there are endless details which could be argued, this is a significant methodological and conceptual contribution of wide applicability. The author states his purpose "to examine, with reference to Jamaica, how best to measure changes in levels of living over time; to examine the adequacy of data available in Jamaica for this purpose; and incidentally to put together such relevant series as I could obtain."

A field study was not attempted; aid was received from the Central Bureau of Statistics in Jamaica, the Institute of Social and Economic Research at the University of the Central West Indies, as well as from specialists both within and without the Government and the University. Moser's return journey was routed via New York where he discussed problems related to his undertaking with experts in the Statistical Office of the United Nations.

Interest of statisticians centers largely in parts II and III, the latter dealing with "The Main Methodological Issues." A number of statistical series (the earliest be-

[1] Report of a Committee of Experts convened by the Secretary-General of the United Nations jointly with the I.L.O. and the U.N.E.S.C.O., "Report on International Definition and Measurement of Standards and Levels of Living," *Document E/CN.3/179 and E/CN.5/290*, New York, 1954, hereinafter referred to as the U.N. Report. For F.A.O. comment see *Document E/CN.3/179. Add. 1*, 18 March 1954; for W.H.O. comment *Document E/CN. 3/179. Add.2*, 5 April 1954. See also ch. VII, "Social Action to Improve Levels of Living" in Robert E. Asher, Kotschnig, and Brown, *The United Nations and Economic and Social Cooperation*, Wash. D. C., Brookings Institution, 1957.

ginning with 1879 and some shown separately for the 14 parishes into which the island is divided), some never previously presented, are given in some 50 tables in part II, "Case Studies: 1. Education; 2. Nutrition; 3. Health; 4. Housing; 5. Income, Expenditure and Consumption; 6. Concluding Remarks." This part treats with nice discernment data limitations and problems of interpretation. Part I deals with "Concepts and General Approach" and part IV with "Future Work." Anyone concerned with field work in or theoretical concepts of consumption economics should read the entire study, as it is difficult to do justice in a summary.

Points of departure are the definitions and method suggested in the U. N. Report. These developed out of the problems of defining and measuring "levels of living (what is)" as distinguished from "standards (what might be)" or "norms (what ought to be)" and of making international as well as intertemporal comparisons of such levels.[2] The level (or levels) is considered as being made up of a number of components such as health, food and nutrition, education, conditions of work, and so forth through a list of twelve, and the measurement of each is first tackled individually, through various indicators. (The thorny matter of their combination into an over-all gauge is discussed in part III.)

In parts II, III and IV he demonstrates the possibility of and necessity for comparing actual levels with standards (subjectively desired or perhaps based on achievement elsewhere) or norms (physiologically determined or targets on other bases which may be set by international agencies or national governments). He takes account of the interrelations between various components of levels of living and the interplay between levels and standards over time as well as in place comparisons.

As crucial measures he seeks indicators, where possible classified geographically and by economic groupings (to yield distributions about the average) of (1) social resources, both physical and human (e.g. schools, hospitals, teachers, doctors) as well as of (2) their utilization (e.g. school attendance) and of (3) the end-product toward which they are aimed (educational level attained, adequate health).

He suggests there is no point (no more than to combine capital and income) in combining indicators relating to resources with those relating to utilization or end results. For construction of composite indexes (separately) for resources, utilization and end products, he reviews the theoretically possible weighting alternatives. He suggests (p. 86) as the most reasonable basis "factor analysis" as used in American studies on farm level of living indexes, with consideration also of a technique of "discriminatory analysis" (p. 65).

A quite different kind of combination of indicators of the first type more akin to national income accounting than to construction of index numbers is sketched (pp. 87–8) as a perhaps more useful new tool. At regular intervals a "social resources account" (human as well as non-human), analogous to a national "balance sheet" could be drawn up for a country. Moser would reserve a place of priority for that "common measure of economic development . . . national income per head (at constant prices), but to get any idea of resultant improvements in social conditions and welfare one should surely go beyond it. This is the function performed by the many indicators proposed in the U. N. Report and the few I have covered in the present study. I now suggest that it might be worth trying to combine the resources indicators into a social resources account and, less hopefully, to assess the return obtained on these resources in terms of improved social conditions" (p. 88).

---

[2] Moser is concerned that possible confusion for a time may result from the agreement on the term "level of living" for what in British usage has commonly been called "standard of living," but he follows the terminology recommended. He notes that at the 1956 meeting of the U.N. Statistical Commission "conditions of living" was suggested in place of "levels of living." I am inclined to agree with him that this seems well worth considering.

Although he leaves many unresolved questions for further research, there is a nice, orderly marshaling of issues. Though he eschews specific recommendations, he cites many possible advances in administrative statistics at negligible or modest cost, and ways of enhancing partial surveys already programmed.

He urges more intensive and informal field work than associated with modern questionnaire surveys, ideally "by a team of experts—e.g. sociologist-statistician, nutritionist, doctor—aided perhaps by one or two high grade fieldworkers, who should be people experienced in social work or something similar." He suggests that each household be studied over a period of a year to see conditions at their best and worst. "Allowing for all the preparatory work, and analysis and reporting, a total period of three years would probably be necessary." A detailed plan along these lines, linked with nutrition and preventive medicine programs might interest governments, research committees, foundations and the United Nations, particularly the International Labour Organization.

Statistical Sources for Market Research. *The Market Research Society.* Lingfield, Surrey, England: Oakwood Press, 1957. Pp. vi, 32. 10s 6d. $1.60.

The Sources and Nature of the Statistics of the United Kingdom, Volume 2. Edited for the *Council of the Royal Statistical Society* by Maurice G. Kendall. London: Oliver and Boyd, 1957. Pp. viii, 343. 30s.

PETER B. KENEN, *Columbia University*

THESE two guides to British statistics are happy complements. The first is designed to lead those of us who may have need of certain British statistics to the series we require. The second is designed to protect us against error in the use and interpretation of those statistics.

The slender volume published by the Market Research Society is a classified list of the business statistics of the United Kingdom. It indicates the frequency, the form, and the place in which they are published and occasionally tells us something about the way in which they have been collected. It does not pretend to be comprehensive; the Society apparently takes a narrow view of market research, for it neglects entirely statistics on prices and on the banking system and the capital markets. What is more, it makes no attempt to appraise or compare the series to which it refers. But it does describe the principal sources in each of the areas it covers and it refers with pleasing frequency to the trade association and other nongovernmental data which the novice will not often discover on his own. It is indexed.

The second of these two books is a collection of specialized monographs. These were first published in the *Journal of the Royal Statistical Society* but have in many cases been revised for inclusion in this volume. Nineteen essays describe and critically appraise the available data on population; on the use of land; on the national income; on local government and regional economics; on banking, insurance and financial investment; on advertising; on medicine, medical services, education and crime; on international air transport and telecommunications; and on the clothing, food, petroleum, timber, and Scotch whiskey industries. They thereby fill many of the gaps left by the first volume in the Royal Statistical Society's series.

There are, however, three areas which deserve but have not yet received systematic treatment. Neither of these volumes thoroughly examines the available statistics on prices, on central government finances and the public debt, or on the balance of payments. These subjects are surveyed elsewhere[1] but ought to have been included in this definitive study. They require and have not yet received the sophisticated treat-

[1] See, e.g., Ely Devons, *An Introduction to British Economic Statistics.* Cambridge: Cambridge University Press, 1956, and Edward Nevin, *The Problem of the National Debt,* Cardiff: University of Wales Press, 1954.

ment that characterizes this collection. There are, moreover, some serious gaps within certain of the articles in the Royal Statistical Society's compendium. W. M. Dacey provides a superb analysis of the available data on Britain's clearing (commercial) banks. Not only does he analyze the published statistics, he often adds to the data at our disposal. But he neglects entirely the discount market and the merchant banks. He also passes over the problems produced by recent changes in the clearing banks' methods of valuation. John I. Mason and David Sachs, examining the statistics on financial investment, make no mention at all of *The Economist's* excellent index of securities prices, nor do they mention the official statistics on new capital issues. And A. L. King, writing on petroleum statistics, might have said something about the finances of the industry, a subject of considerable interest to students of Britain's balance of payments position.

A few of the articles in this volume fall short of the high standard maintained through most of the book. J. E. G. Utting's analysis of the national income statistics, for example, tends to degenerate into a catalog of the contents of the National Income *Blue Book*. He seems reluctant to comment upon the preparation or reliability of the data. Bernard Benjamin's discussion of medical statistics is painfully complex. It is much more of an agenda for research than a survey of the materials available. A reader not already acquainted with the data he considers cannot expect to profit from a study of his article.

But these blemishes do not much detract from the contribution which this volume and its predecessor make to our knowledge. They comprise a thorough and critical guide to the published statistics of the United Kingdom. The Editor and his associates are surely to be congratulated. If they would provide an index to their compendium we should be truly indebted to them.

The Economics of the Fur Industry. *Victor R. Fuchs.* New York: Columbia University Press, 1957. Pp. x, 168. $5.00.

G. R. GREGORY, *University of Michigan*

ALTHOUGH the title suggests a work of considerably greater scope, this study was pointed toward a dual objective: (1) to find out what makes the fur industry "tick" the way it does, and (2) to throw some light on what the author feels are the two outstanding problems in the industry—short-run instability and long-run stagnation and decline. These more limited objectives are achieved: Moreover, they are achieved in a most readable fashion.

To reach his first goal, Fuchs includes a concise but reasonably detailed description of the fur industry. Industry is interpreted broadly, including the production of raw fur skins on the one end and the retailing of fur garments at the other. The fact that uniform treatment is not given all segments of the industry proves to be no drawback since it is the little-known aspects of the industry that have been emphasized. Fur manufacturing, for example, is described in considerably more detail than is the retailing of fur garments.

In accomplishing his second objective—treating the problems of short-run instability and of long-run decline—Fuchs examines monthly time series from 1946 through 1952 for evidence of nonseasonal cyclical movements (seasonal variation is obviously important and adjustments are made in the data to eliminate this effect). It is readily apparent that non-seasonal fluctuation in the fur industry is "fairly drastic." It is less clear that 1946 through 1952 provides sufficient evidence to make meaningful the statement, "Fluctuations within the fur industry follow a discernible internal pattern with respect to amplitude, timing, and the transmission of impulses from one stage to another." The conclusion that expectations play a leading role in causing the fluctuations, while not surprising, seems well founded.

The section on consumer demand, handicapped by a lack of statistical information, leans heavily upon the author's apparent intimate knowledge of the industry. This knowledge is also put to good use in the discussion of the industry's competitive structure. The conclusion that the highly competitive nature of the industry is one of its major handicaps seems well supported. The inclusion of a forthright set of recommendations, while admittedly impossible to carry out because of legal, institutional, and psychological barriers, still provides the industry with a basis for self-improvement.

An outstanding feature of the book, and one which this reviewer would like to see employed more often, is the inclusion of an appendix containing much of the data used in the analysis as well as a description of the derivation of the time series utilized.

The book should be of primary interest to members of the fur industry itself and to economists interested in industrial organization or in the theory of the firm or business cycle theory. It should also be received warmly by students of wildlife management—a field almost notable for its lack of economic study.

Problems in the International Comparison of Economic Accounts. *John W. Kendrick*, Editor. Studies in Income and Wealth, Volume 20, The Conference of Research in Income and Wealth, The National Bureau of Economic Research, New York. Princeton, N. J.: Princeton University Press, 1957. Pp. x, 404. $8.00.

EDGAR O. EDWARDS, *Princeton University*

THE progress made in this book toward the improvement of international comparisons of economic activity is almost as imposing as the conceptual and practical hazards that remain. The subject matter covered in the six contributions ranges from an attempt to distinguish theoretically between economic and noneconomic activity to the construction of a comprehensive system of social accounts. Other contributors deal with index number problems, the reconciliation of different concepts of the government budget, treatments of government activity in social accounts, and international transaction accounts. The critics whose incisive comments follow the various articles are often as thought-provoking as the contributors. Despite the uneven quality that is characteristic of most multiple-author works, the achievements of this volume are substantial.

The most solid advance is scored by Morris Copeland, who constructs a comprehensive system of social accounts in which he integrates domestic product and income accounts with balance sheets and flow-of-funds statements. Input-output data, requiring sectoring by establishment rather than by ownership, he would report in value terms in supplementary tables. The care with which he has related his system to the recently proposed UN system is itself no small accomplishment.[1] Too often such proposals are presented devoid of roots.

The adoption of Copeland's system of national accounts would involve substantial changes in the data presently reported for the United States. It would entail among other things (1) a switch from the national income concept to the domestic product concept (a move which Ruggles supports but which Jaszi questions); (2) a private noncorporate account which would, in common with the government account, include both current and capital transactions on the grounds that decision-making is better analyzed in this way; (3) an account for existing fixed asset transactions and a contra entry in the private noncorporate account which would convert private noncorporate saving into net lending—a change that would permit the establishment of an account for financial transactions and the abolition of the saving and investment account. Most of these suggestions have as an aim the improvement of behavioral

[1] *A System of National Accounts and Supporting Tables*, United Nations, Statistical Office, Studies in Method Series F, No. 2, 1953. In the process Copeland subjects this document to the most searching review I have seen.

analysis. This is an important objective in its own right, but it must be conceded, I think, that its achievement rests on the intimacy with which the data on a country are related to its own institutions. It follows that the standardization of data and precision in behavioral analysis are likely to be incompatible goals. The purposes to be achieved by standardization involve instead welfare-rooted comparisons, a matter that is treated gingerly, when it is treated at all, not only by Copeland but by other contributors as well.

To improve comparability between industrialized and underdeveloped areas, Copeland approves of more imputations and grants a temporary need for a rural sector. His standardized system emerges, however, heavily oriented toward industrialized countries on the rather optimistic assumption that "nonmoney communities will presently cease to exist." Leontief in his comments advances the interesting hypothesis that much of the difficulty in getting comparable data on underdeveloped countries stems from just this kind of monetary premise and the valuation problems it entails, rather than from a shortage of raw quantity data. The use of quantity data, he feels, should be further explored as a basis for a system of interlocking accounts before adopting a comprehensive system based on monetary values.

Gerhard Colm in his article reviews some of his earlier positions on the treatment of government product in the national accounts. He continues to feel that any attempt to identify government intermediate product would give rise to larger errors than now exist in the United States data. On the other hand he has come to the conclusion that the imputation of interest on government capital would improve the accounting totals. Fabricant takes convincing exception to the first view arguing that if *any* amount of government product can be identified as intermediate the resulting estimates of gross national product would be improved.

In developing the second view Colm suggests that the inclusion of *imputed* interest in national product would obviate his earlier proposal to include some of the *actual* interest payments of public bodies. The mingling of these two problems for whatever reason seems unjustified to this reviewer. The imputation of a return to capital remains a problem whether or not the money used for its construction is borrowed. Similarly the treatment of actual interest payments should be viewed independently of the uses to which the borrowed funds are put. The product for which interest is paid is the current use of purchasing power. One must examine the use to which the funds are put only to classify this product as final or intermediate. It might be decided, of course, to exclude the final product from the accounts on the ground that it arises from a redistribution of other more tangible products. But this argument is quite different from a denial that there is any "*current* counterpart" to the interest payment which is Colm's position.

A useful and detailed reconciliation of the conventional, cash and Commerce statements of government receipts and expenditures is provided by Marilyn Young. Her contribution, a valuable and impressive reference document, is supplemented by the comments of Block and Landau of the United Nations who summarize some of the budget techniques used in other countries. The anachronistic nature of the United States treatment of government capital accounts is glaringly exposed.

Working on the fringe of established systems of data, Herbert Wooley discusses the problems involved in elaborating a system of international transaction accounts. He is concerned mainly with how best to group nations (what purpose should dominate?) and how to handle multi-country petroleum, shipping, and merchanting transactions. Some of his transaction problems might be simplified if one could build on country accounts, with which the international ones should mesh, that are themselves standardized whether on a national product or a domestic product basis.

To those who use index numbers without seriously questioning the fundamental assumptions, the Brady and Hurwitz article should prove refreshing. The recent

multiplication of products, product qualities, and attached services (guarantees, for example) leads them to question the validity of matching one product with another for the construction of intercountry price indices. To avoid "the need for imputing identity to particular commodities," they suggest "matching income levels at which the same percentage of income was spent on food, [necessities or other standards of equivalence]." Two groups of people that match in the selected standard are presumably equally well off. If the hypothesis is accepted that the proportion of income spent by different groups within a country is a function of *relative* incomes, this technique might demonstrate, as Tobin suggests, that real income is the same all over! But given the unsatisfactory results achieved with index numbers, the hypothesis certainly warrants further testing.

Starting from the "changelessness of man as a social being" and the universal "antithesis between ends and the limited means" Kravis searches for a means of delimiting economic activity. He proposes "(1) the rule of remunerated activities" and "(2) the rule of sensitivity to rewards." The sensitivity rule, which receives the brunt of Hagen's and Viner's criticisms, is essentially a cross elasticity—the elasticity of time supplied to a nonmarket activity relative to the price of time supplied to a market activity. Thus if a rise in wages induces a man to spend less time in his garden relative to his other nonmarket activities (say, attending church), the garden work is an economic activity. But every activity, market or not, is only partly economic, and the economic-noneconomic ratio for any activity (say, garden work) undoubtedly varies substantially from one individual to another; where one decides "economic" stops and "noneconomic" begins can be likened to defining different shades of grey as either black or white. Yet the criterion, sharpened, may aid in identifying the extremes.

The varied subject matter in the book would make it difficult for the average reader to use profitably were it not for the carefully prepared subject index, and the excellent summary written by John Kendrick as an introduction to the volume. All the contributors, experts in this field, would concur with his view that "comparisons of the national economic accounts can seldom produce wholly unambiguous meanings." Can we hope that "standardization," toward which this volume is a successful step, will not lead the "amateurs" to ignore the warning?

Demographic Analysis: Selected Readings. Joseph J. Spengler and Otis Dudley Duncan, Editors. Glencoe: The Free Press, 1956. Pp. xiii, 819. $9.50.

GEORGE J. STOLNITZ, *Indiana University*

THIS collection of sixty-three papers is a companion volume to the editors' *Population Theory and Policy*, also published in 1956. The basis of division between the two readers is far from clear-cut but roughly the latter is oriented about the theoretical or welfare aspects, in general terms, of population analysis. The editors have sought to provide the foundations of such analysis in the present book, by illustrating demographic thinking in relatively "concrete" or "empirical" problems. The papers, all previously published, are arranged in seven groups: (I) Past and prospective growth and distribution of world population; (II) Mortality; (III) Fertility, nuptiality, and replacement; (IV) International distribution of population and migration; (V) Internal distribution and migration; (VI) Population composition and utilization of human resources (including quality of population); and (VII) Selected regional studies. Contained in the last are descriptions, variously approached and for varying dates, of France, Italy, Germany, the Soviet Union, Brazil, Puerto Rico, Africa, Egypt, China, Japan, Indonesia, Southeast Asia and Burma. Brief introductions are given before each group of papers and a bibliography, taken largely from the postwar journal literature, is provided at the end.

Alone or in combination, these readers are ambitious ventures by any standards;

they are easily the most ambitious ever undertaken in the field of population. The editors have aimed at an *Analysis* which "may be used by itself as a relatively self-contained basis for a course in demographic analysis, in combination with the companion volume for a course emphasizing both theory and research, or as a sourcebook of outside readings in conjunction with one of the standard texts on population problems or methods of demographic research." An apparent further aim, not stated explicitly, has been to accommodate the nonstudent seeking a convenient introduction to the field. The editors have made a welcome attempt to overcome the continuing need for materials suitable to either purpose.

How well they have succeeded in their "course" objectives will obviously depend upon teacher and class, but it seems safe to forecast that it will be the teacher rather than the student who will benefit in the main. Papers written by specialists for their peers are likely to presuppose context and the student will have to seek elsewhere for perspective and a sense of proportion. Although the editors' introductions are generally excellent, they are probably too concise and allusive to be very profitable to those who need them most. The decision to pay little attention to techniques is understandable but will be considered by some to be unfortunate, considering the returns to be expected from even marginal efforts. The only current source of general instructional interest which also provides an introduction to techniques is the text by Landis and Hatt, which is, however, almost exclusively concerned with American population.

Pedagogical issues aside, both the specialist and nonspecialist should find much in these pages to reward their interest. As the above titles suggest, the range of topics is so broad that it may seem ungracious to ask for more. Yet there are some surprising gaps in emphasis. The remarkable mortality declines now under way in many underdeveloped areas—probably the most important single event in international demography since the war—receives only slight mention. The uninitiated reader is likely to come away with impressions on this subject which would be appropriate to the 1930's. The postwar recovery of fertility in Western nations is similarly neglected; the prevailing tone of most selections is more nearly that of 1940 than 1955. A decade of developmental efforts in the underdeveloped areas, in the face of accelerating population growth, finds no representation at all in these papers. The informational revolution stemming from the introduction of sampling methods in many parts of the world is overlooked. Paradoxically, many of the papers which seem most impressive in retrospect are in the sections on population quality and migration, although these areas have not been in the forefront of demographic investigation in recent decades.

The statistician will find little of explicit technical interest in these pages but will surely be struck by the wealth of problems to which his skills could be profitably applied. To name a few: such questions as the relations between economic and fertility fluctuations among populations practicing birth control, proper methods of standardization under incomplete controls for residual heterogeneity, the effects of changing composition on vital trends and differentials among population aggregates, the stochastic and systematic factors relating cross-sectional and longitudinal data, an appropriate identification theory for distinguishing separate and composite effects on fertility, or the selective characteristics of migrant groups, have yet to register the impact of modern statistical approaches.

The 1953 Pilot Population Census for the First Population Census in Sudan. Prepared by *K. J. Krótki*. Khartoum: Sudan Government, Ministry for Social Affairs, Department of Statistics, 1955. Pp. 183. 1£ Egyptian or 1£ 0s 6d English currency. Paper.

W. PARKER MAULDIN, *The Population Council*

THIS beautifully printed book is a detailed report on the Pilot Census in Sudan in which slightly more than 52,000 persons in 10,511 households, in 9 sample

areas were enumerated. The detailed table of contents of almost 8 pages is indicative of the care with which this report has been compiled. These initial favorable impressions are strengthened by the presence of a map, photographs of local scenes, and of enumerators in various settings, a copy of the English and Arabic questionnaires or schedules used in the enumeration, etc. Also, the non-technical parts of the report have been written with an eye to attracting the interest of the reader and a number of examples of the unusual are given. The Pilot Census itself incorporated at least one novel feature in one sample area where a lottery was held for those participating in the Census, and emphasis was placed on the fact that prizes would be given only to those included in the Census and to those whose entries on the Census forms proved, upon investigation, to be correct. Moreover, the author has been quite frank in enumerating the problems encountered, and even in listing the failure to obtain a number of the objectives of the Pilot Census.

The primary objectives of the Pilot Census were: (1) to determine variances in order to compute the sizes of samples necessary in different areas; (2) to test the questionnaire; (3) to determine whether "foolproof plans had been made for the translation, scrutiny and coding of the answers"; (4) to test the form of organization at headquarters, including machine analysis and preparation for it; (5) to test the existing Province and District organization; (6) to test the provisions made for carrying out the census in the field; (7) to test the various types of enumerators and supervisors available in a country with areas at such diverse levels of development and sophistication; (8) to determine the number of field personnel required and the time taken by different kinds of enumerators; (9) to test the efficiency and practicability of the proposed enumeration techniques; (10) to provide information on which to base a realistic census budget; (11) to provide information regarding optimum seasons of the year during which enumeration should be carried out; (12) to test the various arrangements for dealing with people speaking about one hundred different languages; (13) to determine the feasibility of using a pre-determined classification of items such as lists of tribes, languages, occupations, etc., and (14) to find out the truth about people's superstitions and objections to being counted.

The sampling plan of the Pilot Census was to obtain a sample from:

11 areas of the country representing different types of settlement, modes of life, etc.
4 Omodias within each area (except in one town which was to be completely enumerated).
16 Sheikhships within each sample area (except in the town).

The sampling fractions for enumeration of households were planned as 1, $\frac{1}{2}$, or $\frac{1}{5}$, depending on the type of living arrangements found within each Sheikhship.

The actual samples drawn deviated from the theoretical plan in several important respects: two areas refused to cooperate; the desired number of sample Omodias was obtained in only half of the 8 areas in which the Omodia was the sampling unit; in only half the areas were as many as 16 Sheikhships included; and the sampling fractions used were quite different from the plan. In some instances the reasons for failure to follow the plan were given, and were understandable, but in other cases the stated reasons were trivial, and in still other cases no reasons were cited. The failure to follow a prearranged plan as closely as possible is important for a number of reasons, not the least of which arises out of the fact that after the samples had been drawn and the field work completed a determination was made as to "What is the exact area of which each sample can be deemed to be representative?" We are told that "no harm was done and perfectly homogeneous and reasonable areas were arrived at, except perhaps in the case of Jebels and Kababish." Unhappily, this part of the report is poorly organized, and it is time-consuming to find the various parts that explain the plan and the results. Details regarding setting boundaries for areas of which the sample was deemed to be representative are not available to the reader.

The basis of sampling within a Sheikhship was (1) to make a rough approximation of total population within the Sheikhship on the basis of available lists (tax, tribute, heads of households, etc.); (2) furnish the enumerator with a list of persons to be enumerated, and (3) instruct the enumerator to enumerate a specified number of households nearest to each of the listed names, or to enumerate all persons within the area chosen for the taking of the subsample. The estimation procedure was based on the assumption that listed and nonlisted households were randomly distributed within the inhabited areas. The validity of this assumption is highly questionable, and unfortunately there is no indication that any attempt was made to determine whether or not the bias in such a procedure was small or seriously large. In areas where the "nomadic technique" was employed, for example, the following instructions were given. "To keep the expenses within limits and to cut down the amount of travelling and walking to reasonable proportions, the general instructions to enumerators were to go to the largest congregation of members of the given Sheikhship, usually at a watering place. There he was to enumerate as many followers of the given Sheikh as he could lay hands on." The author states that "omissions from the tax list are random," but offers no evidence to support this assumption. If the Sheikhship extends beyond the watering place at which enumeration was carried out, there is the further assumption that the ratio of non-listed households to listed households is the same for all other areas in the Sheikhship as that found at the place of enumeration. This procedure does not live up to the author's claims for sampling: "Where people lead a nomadic mode of life, complete coverage could never be guaranteed and double counting could not be avoided during a full count census. With a sample census these problems can be dealt with satisfactorily."

But in spite of serious deficiencies of the types mentioned above, this may be a useful guide for those working in underdeveloped areas such as the Sudan. The report is quite frank in listing its "failures," such as: unsatisfactory classification of tribes, languages, and occupation; procedure for listing heads of households; the post-enumeration survey which was not well planned, and did not produce satisfactory results; occasional failure to ask about temporary absentees; translators not being consistent in translating some words; etc. It is questionable whether or not the cost figures shown in this report would be helpful in another country, but certainly they will be of use in the Sudan. The discussion of nonsampling errors is perfunctory and unenlightening.

If one attempts to assess the accomplishments of the Pilot Census relative to the stated aims, the resulting score sheet may be somewhat discouraging. In part this is due to the fact that the aims were more extensive than they should have been. In spite of the deficiencies of the study, examples of which have been illustrated above, The Pilot Census yielded much valuable information for the Sudan, and, in addition, the description of the practical problems and solutions attempted will be of considerable value to persons attempting sample surveys or censuses in similar areas.

Information and Control. *Leon Brillouin, Colin Cherry, and Peter Elias,* Editors. Volume 1, No. 1, September, 1957. New York: Academic Press, Inc. Pp. ii, 89. $10.00 annual subscription for 4 issues.

ACCORDING to its own statement, *Information and Control* proposes "to publish, papers which make significant contributions to the theories of communication, computers, and automatic control, and also papers which present experimental evidence or theoretical results bearing on the use of ideas from such theories in any field to which the ideas are relevant . . . it will from time to time publish editorials, or papers of an editorial nature, discussing views of the relationships between fields . . . the editorial board is international in character.

The first number contains nine short articles, six of which are by United States

authors and the others by French, Swiss, and English, but all are written in English. The longer articles have short abstracts.

D. D. F.

Operations Research for Management, Volume II. *Joseph F. McCloskey and John M. Coppinger*, Editors. Baltimore: The Johns Hopkins Press, 1956. Pp. xxxvi, 563. $8.00.

W. W. COOPER, *Carnegie Institute of Technology*

THIS book is a collection of writings by persons active in OR—operations research—and related areas. Prior to their selection, all papers were presented at The Johns Hopkins University Informal Seminar on Operations Research during 1953–54 and 1954–55. The book is designed as a series of vignettes. Subtitled "Case Histories, Methods, Information Handling," it attempts to provide material for a picture which will enable "management" to read meaning into the multifarious forms which operations research might assume.

Three common purposes which such "collected writings" might serve in the literature of science may usefully be noted. First, as in the case of the A. Wald or R. A. Fisher collections, this device might be used to bring together a series of contributions from one person (or a team) so that they might be better understood and evaluated as a coherent whole. Second, more or less distinguished contributions by independent authors might be organized under one topic or theme and thereby demarcate some field of endeavor and supply an advanced starting point for those persons who wish (a) to obtain access to these results for their own work without being required to trace a labyrinth of specialized literature or (b) to attain a jumping off point at the frontiers of a series of developments. The Cowles Commission Monograph, *Statistical Inference in Dynamic Economic Models*, edited by T. C. Koopmans is an example.

The (perhaps) extreme examples thus cited, with their emphasis on outstanding research problems, have been deliberately selected to distinguish another purpose which such a collection might serve: explication for a lay audience which either will not or cannot enter the professional literature. It is this third category which the present book attempts to serve (at least by title and stated intent).

OR is a newly developed field. There is as yet no consensus on its form, content or specific areas of application. This may have tempted the editors to enter the second (as well as the third) category of purposes. An indication of this is apparent in the changed emphasis in the editors' introduction to this volume—the second to appear in a series of three promised volumes under the same title. The extension was probably also unavoidable (a) because of the rather considerable progress in operations research between the appearance, in 1954, of the first volume and manufacture of the second and (b) because the presentations incoporated in this volume were intended for a seminar with a rather healthy representation of experts. The book is as much an outgrowth of these forces as of a process of editorial selection and pruning. The resulting mixture makes it necessary to balance and assess the accomplishments of explication against the quite different requirements of technical excellence and completeness. Both are attenuated.

Nevertheless, the book can be recommended—even recommended highly—to the following two kinds of persons: (1) those who might want "to buy" operations research and (2) those who might want "to do" it—or find out whether they have been doing it. This recommendation carries with it a reading prescription. Both groups should read the book rather rapidly. This will provide perspective on what OR is about and, more particularly, the nature of the problems it has dealt with, its philosophy, and the tools it has utilized. (They may also find that they are less likely to draw a picture of what OR might be able to do—e.g., in as yet untried fields and problems—and where it is likely to go in the future.)

With few exceptions (e.g., Merrill Flood's piece on "The travelling salesman problem")[1] those who treat this as a category 2 book will be disappointed, especially if they believe that careful study of its contents will provide an advanced starting point from which they may either (a) adapt OR "findings" for use in their own field of interest or (b) jump to the frontiers of operations research. Such persons had better look to the professional journals, such as *Operations Research* or *Management Science*. Fortunately, these journals are recent in origin so the labyrinth is rather easily traversed. If some kind of preliminary "rounded" introduction is also desired, these persons will find a better book in *Introduction to Operations Research* (New York: John Wiley and Sons, Inc., 1957), a one-volume compilation edited by C. W. Churchman, R. L. Ackoff and E. L. Arnoff.

The merit of the present volume is for those who seek a rapid survey. They will find the book divided into four main parts of uneven length and emphasis: (1) Introduction, (2) Case Histories, (3) Methods, and (4) Information Handling (small group psychology). Much, but not all, of this material has been published elsewhere, especially in *Operations Research*.

The *Introduction*, consisting of an article "From Plato to the linear program" by the Earl of Halsbury needs to be read twice in order to allow hackles raised on a first reading to settle on the second. One way of regarding this piece is as an illustration of an often encountered point of view which divides (presumably) thoughtful and equally well educated people into two groups: (1) those who are "narrow technicians" and (2) those who are not. Apparently, the Earl of Halsbury, or at least this article, belongs in the latter category. Certainly it is sweeping in scope. But almost every "technical" point—the ventured interpretation of Plato, animal and human ecology, the characterizations of J. M. Keynes, and an "application" of linear programming—may, at best, be characterized as a whopping half truth and, at worst, as misleading, ill-conceived, or erroneous. Nevertheless, the main conclusions are correct (or at least compelling) and display a grasp of the nature (and history) of scientific evolution. The point is made that OR is only a recent manifestation of the grand sweep of this evolution from ancient to modern times. Improvement in methodology is gradually (and in accelerating fashion) allowing science to break from its laboratory confines and confront those tasks, in the world of practical affairs, which the classic philosophers originally conceived as objects of rational inquiry. Those besetting problems of society which have not yet yielded to any other approach (force, philosophy and religious belief) will, it is hoped, prove more tractable to these emergent methods of science. It should be noted, however, that the scientific task as thus conceived involves more than acquiring substantive knowledge. It also requires the evolution of adequate means for dealing with the huge scope and size of such problems (already apparent in many OR areas) involving thousands and hundreds of thousands of variables, equations, and related parameters.

Part 2, *Case Histories*, is a potpourri, as it should be. It ranges over applications in a variety of fields: military, including an historical resume (by Leroy Brothers), industrial (e.g., Dunlap and Jacobs' analysis of dragline operations in mining phosphorous), and local government (e.g., L. C. Edie's analysis of traffic delays for the Port of New York Authority). These applications are supplemented by further proposed applications in each area. The latter include the reported studies by Roger Crane (Analysis of a Railroad Classification Yard), William Vickrey (Revising New York's Subway Fare Structure)[2] and Stanley Davis (A Study of Combat Stress in

[1] The articles by Christie and Luce form another exception.

[2] Readers should be warned not to confuse Vickrey's equations or methods of estimation with those that might be founded on statistical methodology. Joint-cost problems, related traffic demands, and attendant difficulties apparently led Vickrey to an informal "judgment-accounting" process for deriving his parameter estimates and to an informal model (not completely specified in the article) for interpreting his results. The "team approach" is often praised in operations research as a device (a) for bringing the talents of experts in different disciplines to bear

Korea). The tools utilized include statistics (with relatively little emphasis on sampling and experimental design), probability (including stochastic processes), mathematics, and a variety of simulation and approximation techniques. The concepts are also drawn from a variety of fields—business, economics, physics, etc.

The next part of the book, *Methodology*, again covers a range of topics. Some parts are simply attempts, at an elementary level, to acquaint OR workers with available methods. W. J. Youden's exposition of experimental design is an example. Others represent responses to the practical need for methods of solutions. As in Glen Camp's "Bounding the solutions of practical queuing problems," these are attempts to deal with a class of problems whether encountered in OR or elsewhere. Still others are direct responses to situations confronted in an OR context. R. Zimmerman's "Monte Carlo model for military analysis" and Walter Cushen's "Operational gaming in industry" are cases in point. Finally, some attempt is made to balance the preceding materials by articles pointed toward new areas and possible problems of the future. Thus, D. B. Hertz and G. J. Feeney jointly essay a statement of a formal model for an OR application to the structural design of managed organizations and C. W. Churchman poses the question of whether OR practitioners (collectively and individually) should accept the fundamental value judgments of their clients, or whether they should subject these judgments to OR inquiry. Neither paper gets very far. But the Pandora's box that Churchman opens cannot be closed without disappointing, in important ways, the levy that the Earl of Halsbury would ultimately lay on scientific inquiry.

The final section of the book, *Information Handling in Organized Groups*, consists of a series of reports by L. S. Christie and Duncan Luce. Much of the material rests on work done at the M.I.T. Research Laboratory of Electronics and reported elsewhere—e.g., in the M.I.T. Reports (1952) "Communication and learning in task oriented groups" by Christie, Luce and J. S. Macy, Jr., and (1954) "Information flow in task oriented groups" by Christie, Luce and D. H. Hay. Some updating is added, however, including discussion of findings by others such as Guetzkow and Simon. Emphasis is on laboratory experiments where small groups are allowed to communicate with each other over certain network channels and are given certain simple tasks to perform. Assessments are made of efficiency (time to perform task) learning (improvement in repeated trials) morale (reaction of group members to tasks and organization) and other psychological factors. Originating in the constructs of group psychology (e.g., as envisioned by the late Kurt Lewin and, more immediately relevant to the reported studies, the suggestions of Alex Bavelas and Harold Leavitt) these studies extend into other areas as well: mathematics (topological analysis of graphs and networks), statistics (experimental design), machine simulation and analysis, information theory extended to semantic content, etc.

The reports in this final section are highly suggestive and appear to contain many promising insights into the area of organizational analysis and prescription—an area which OR is now only beginning to plumb. The studies are not yet operations research, however, and will not become so until means are found for carrying these insights beyond the laboratory. While he should be prepared to use new scientific knowledge (from a variety of fields) as it becomes available, the OR practitioner is also finding that he must be equally prepared to go ahead of this knowledge as his problems and opportunities require. Indeed, as he does this he can provide a service to science as well as to management and thereby help to better the state of both. The area covered by Christie and Luce may well become a case in point judging from recent trends in operations research.

---

on particular problems and (b) for stimulating research in these disciplines. The difficult nature of the statistical (and mathematical) problems encountered in this study might naturally have suggested this kind of approach to Vickrey but there is no report of any consideration having been given to it.

A significant fact in the growth of operations research is that it is providing a body of practitioners devoted to the entire range of management problems and possessed of the ability to use analytical concepts and disciplined scientific methods in attacking these problems. Statistical methodology stands in the mainstream of the potential developments which may emerge from this growth. To those who desire it, this book provides an opportunity to acquire some "feel" for the course that such developments might take.

The Rationalization of Industrial Decision Processes. Engineering Experiment Station Bulletin No. 163. *William T. Morris.* Columbus, Ohio: State University College of Engineering, 1957. Pp. iv, 117. $2.00. Paper.

CHARLES C. HOLT, *Carnegie Institute of Technology*

THIS bulletin, originally prepared as a Ph.D. thesis, is concerned exclusively with normative theories of decision making, particularly those that might be applied to "engineering" or "industrial" problems. Morris is critical of the older literature in the field of engineering economics which does not get far beyond interest calculations. He sees great hope, as does the reviewer, in the new work being done on decision theory and his book is an effort to make this theory available to engineers.

The chapters deal with: introduction and history, rationality for engineering decisions, the measurement of values, theorems for probabilistic decisions, decision making with incomplete information, judgement, and the timing of decisions.

The "axiom of long-run success" the maximization of expected utility is presented as an "axiom which appeals to the engineer." This version of rationality serves as the basic orientation of the book. Decision problems in which there is an absence of knowledge concerning one or more of the following: utilities, probabilities, possible states of the world, or possible action alternatives, are transformed through observation and learning until they are amenable to rational solution under the axiom of long-run success.

The courage to undertake such a task in 112 pages can only be admired, but the result is a rather cryptic, eclectic, and largely mathematical exposition. Because the exposition is abstract with few illustrative applications, the book is not likely to serve well as an introduction to practicing engineers who are unfamiliar with the field. An exception is the chapter on "Judgement in engineering decisions" which should be immediately useful.

For statisticians the book may be of interest for its broad coverage in a few pages and its references to the literature.

The Computing Laboratory in the University. *Preston C. Hammer,* Editor. Madison: University of Wisconsin Press, 1957. Pp. xv, 235. $6.50.

FRANZ L. ALT, *National Bureau of Standards*

THIS volume contains the papers presented at a conference held at the University of Wisconsin in 1955 for the purpose of discussing problems encountered in establishing computing laboratories in universities. Since electronic computers have come into more widespread use, a number of universities plan to install computers as an aid to teaching and research. The articles presented here illuminate many of the results and obstacles to be expected.

The book begins with three general papers, by C. A. Elvehjem, J. H. Curtiss, and J. W. Forrester, which should be profitable reading for the university president or provost concerned with over-all planning of a computer installation. "*Don't* expect the laboratory to pay its way through computing services to industry. *Do* attach it to

the Department of Mathematics, but *do* make the computer available to members of other departments without inconvenience or cost. *Don't* neglect applications in the social and biological sciences. *Do* acquire large digital machines before analogue ones." Subjects of this nature are discussed, though not, of course, in these words.

Ten papers about applications of automatic computation in science and industry barely scratch the surface of the field. Some of the authors concentrate on the problems solved by them or their laboratories, rather than surveying all computer applications in a given field; and some important applications have been allowed to fall between two chairs. We miss, for instance, descriptions of crystallographic X-ray diffraction analyses, of radiation penetration (shielding), of optical ray tracing, of elastic or molecular vibrations, all of which are bywords in automatic computation. We are grateful, on the other hand, for thorough surveys of computing in astronomy (W. J. Eckert), meteorology (P. D. Thompson), physical chemistry (J. O. Hirschfelder), and for a detailed enumeration of computer applications in the aircraft industry (H. S. Wolanski). A more complete survey of computer applications would be useful to prospective founders or managers of university computing laboratories, since a main mission of such a laboratory is to educate other university departments in the use of computers.

Four papers are devoted to "Future demands for personnel in computing," a question which hardly needed a more elaborate treatment than this. The following group of papers on "Curriculums for computing" are what the department chairman really needs. In two detailed and factual papers by G. E. Forsythe and V. C. Rideout, and two briefer ones by C. W. Adams and D. M. Young, he is given not merely lists of courses but advice on the activities and general attitudes which will promote the ultimate purposes of his computing laboratory.

There are five papers on "Equipping a university computing laboratory," whose competent authors find more than enough space for saying all there is to be said on this subject, and in addition branching out into the philosophy of the laboratory, occasionally with contradictory conclusions, as is only fair. "Organizing and financing" is similarly treated by four writers, each of whom principally describes his home laboratory.

The last chapter of the book, under the heading "Computation in perspective" contains a single paper by J. H. Van Vleck, entitled "Dangerous Gulfs: Some Reflections on the Social Implications of Computing Machines." It presents a deepgoing analysis of scientific-engineering, economic, sociological and cultural aspects far transcending computation. The time for reading these few pages is well spent; the reader is rewarded with startling insights into some of the present trends in our civilization.

Work Sampling. Second Edition. *Ralph M. Barnes.* New York: John Wiley and Sons, Inc., 1957. Pp. x, 283. $7.95.

W. Grant Ireson, *Stanford University*

Ralph M. Barnes has assembled in this small book a working manual for those who have need of sampling techniques to obtain reliable information about working time, performance rating, indirect incentive measurement, demand, and allowances. The book is divided into three parts. The first part is devoted to the history, fundamental statistical techniques, and procedures for making a work sampling study. Part II presents sixteen applications of work sampling studies to a wide variety of problems. Through these applications, Barnes shows that work sampling is a rapid and economical means of obtaining the information necessary for such diverse purposes as setting work standards for indirect labor, improvement of work methods in

nonrepetitive tasks, determination of amount of work to be performed, and the setting of allowances for different kinds of work. He introduces the use of the Waiting-Line theory as an aid in planning for most economical use of personnel and equipment, and shows how the information for the application of these theories can be secured through work sampling. Part III introduces work sampling techniques for performance measurement and rating and the determination of the validity of the results. It also contains a complete reproduction of L. H. C. Tippett's, "A Snap-reading Method of Making Time Studies of Machines and Operatives in Factory Surveys." This was the original paper, published in the Shirley Institute Memoirs in November, 1934, from which the whole field of work sampling or ratio-delay has grown.

The book has been written primarily for the engineer who has little or no statistical background. Chap. 3 introduces the concepts of probability, normal distribution, confidence levels, and absolute error in the most elementary forms with very little explanation of why these concepts are applicable to the subject. The standard deviation of the binomial distribution is presented, cook-book style, as a means of setting the confidence level, and is used to determine the number of observations necessary to obtain the desired confidence level. To save the reader from computing the number of observations for a given confidence level and from the recomputation of the actual confidence level obtained after the observations are completed, Barnes includes in Chap. 5 both tables and charts from which these values can be obtained. He also includes both tables and alignment charts for estimating the absolute error.

The Shewhart control chart concept is explained briefly in Chap. 4, and the p chart is used as a means of testing for homogeneity among subgroups of observations.

The author emphasizes the importance of randomness in the work sampling procedures and provides several pages of random digits. In Chap. 7 he explains in considerable detail how to use the random numbers for randomizing any or all of the factors involved in a work sampling study.

One of the interesting parts of the book is Chap. 22, in which Barnes reports the results of 14 different studies in 8 different companies. The purpose of these studies was to develop a sampling technique that would give an accurate and economical method of determining the performance level of labor activities. Statistical techniques used in the study included the Chi-Squared test for normality of data, Tchebycheff's Inequality, and the Central Limit theorem.

While the book may be of little interest to statisticians, it should be valuable for industrial engineers in two ways. It presents a clear and interesting description of work sampling as a tool. Secondly, it illustrates how statistical techniques can be used by industrial engineers to reduce their work and improve the reliability of their recommendations and conclusions.

Théorie des Fonctions Aléatoires, Applications à divers phénomènes de fluctuation A. Blanc-Lapierre and Robert Fortet, with a chapter on fluid mechanics by J. Kampé de Fériet. Paris: Masson et Cie, 1953. Pp. xvi, 693. 6000 fr in paper; 6500 cloth.

LEO J. TICK, *New York University*

THIS volume, published as part of a "collection of mathematics texts for the use of physicists," represents a complete survey of the probability aspects of random function (stochastic process) theory and applications at the time of publication. Roughly, the book is divided into four parts. The first part contains some introductory material with physical background, and probability foundations. (The last

chapter, really an Appendix, is a brief survey of measure theory, group theory, and others which are used in the text.) The second portion treats of processes with independent increments, including Poisson processes and processes derived from them. This is followed by an extensive study of Markoff chains and processes. The book concludes with a broad treatment of stationary processes.

The use of random processes, as a physical science model, has had tremendous growth in the recent years. Unfortunately, the written material available in English has not been such as to provide a good introduction to an interested novice. It has, mostly, been of two varieties. (1) Those papers or books in which some particular physical problem is posed and solved. The unity of the methods used is not shown and the techniques are not related to any basic probability knowledge a reader might have; often the whole spirit of modern probability theory is missing. (2) Other papers or books written by the purely inclined mathematician in which intensity is purchased at the price of extensity. Very little connection with physical models is displayed. This makes it quite difficult for the physicist or engineer to read. This latter criticism is less true of Markoff and Poisson process work, but it is true about stationary processes. The latter is one of this reviewer's main interests so he should be forgiven for stressing this point.

Taken as a whole this volume is a rich source of material on the theory of random processes viewed from the standpoint of physical applications, at least up to the time of publication. This last point should not be taken too seriously even by a current reader. There really have not been too many advances for the intended audience, i.e. those interested in applications. In fact, as will be pointed out below, there are some topics in stationary process theory which have hardly been treated at all in the English literature.

This book is a fairly long one, and, instead of trying to discuss all of it, it seems, to this reviewer, better to dwell on those aspects which distinguish it, for good or ill, from other works. One of those most striking features is the completeness with which many of the elementary points are treated. This is good, as these are often the parts that the novice finds most difficult, particularly one new to probabilistic thinking. The authors are not adverse to diagrams and writing out each step. In contradistinction, some readers have pointed out that obtuse and difficult points are lightly passed over. Perhaps the physical science audience will not be overly concerned with these deeper aspects of the subject. Another feature which this reviewer found most refreshing was the explicit discussion of vector and multidimensional (time or space parameter) cases. It has always been most frustrating to read mathematical works in which the author states that generalization of two or more dimensions is immediate and obvious. Too often it is neither. This reviewer especially enjoyed the treatment of stationary processes. The introduction of the spectral decomposition through the study of linear filters is especially good. The recent discovery that scattering of radio waves from the troposphere or ionosphere can be a reliable means of over the horizon communication, has created a great interest in multidimensional processes. There is very little in the English probability literature beyond the type of statement given above. However, there are a large number of symmetry conditions possible in higher dimensions and they bear study. The reader will find explicit treatment of these processes in this work in the study of polarized light and turbulence.

This is a probability text; there is no study of statistics except for a few pages on the estimation of spectra. From the above the reader has no doubt concluded that this reviewer commends this volume as an introduction to random functions for the physical scientist or anyone else.

PUBLICATIONS RECEIVED

Bernert, Eleanor. H. *America's Children.* Social Science Reseach Council in cooperation with the Bureau of the Census. New York: John Wiley and Sons, Inc.; London: Chapman and Hall, Ltd., 1958. $6.00.

Beyer, Glenn H. *Housing, a Factual Analysis.* New York: The Macmillan Co., 1958. $6.75.

Beyer, Glenn H., and Rose, J. Hugh. *Farm Housing.* New York: John Wiley and Sons, Inc., 1957. $6.00.

Brown, J. Douglas, and Harvison, Frederick. *High-Talent Manpower for Science and Industry. An appraisal of policy at home and abroad.* Princeton, New Jersey: Industrial Relations Section, Princeton University, 1958. $3.00.

Chamberlin, Edward H. *The Economic Analysis of Labor Union Power.* Washington, D. C.: American Enterprise Association, Inc., 1958. $1.00. Paper.

Davenport, Wilbur B., Jr., and Root, William L. *An Introduction to the Theory of Random Siganls and Noise.* New York: McGraw-Hill Book Co., 1958. $10.00.

Dorfman, Robert, Samuelson, Paul A., and Solow, Robert M. *Linear Programming and Economic Analysis.* New York: McGraw-Hill Book Co., 1958. $10.00.

Goodman, Richard. *Teach Yourself Statistics.* Teach Yourself Series. London: The English Universities Press, Ltd., 1957. 7/6.

Grenander, Ulf, and Szegö, Gabor. *Toeplitz Forms and Their Applications.* California Monographs in Mathematical Sciences. Berkeley and Los Angeles: University of California Press, 1958. $6.00.

Hunt, K. E. *Colonial Agriculture Statistics: The Organisation of Field Work.* Colonial Research Publication No. 22. London: Her Majesty's Stationery Office, 1957. 14s. Paper.

Johnson, N. L., Compiler. *Statistical Exercises.* Part II. Analysis of Variance and Associated Techniques. London: Dept. of Statistics, University College, 1957. 12s. Paper.

Kendall, M. G. *A Course in Multivariate Analysis.* No. 2 of Griffin's Statistical Monographs and Courses. New York: Hafner Publishing Co., 1957. $4.50.

National Bureau of Economic Research, National Accounts Review Committee. *The National Economic Accounts of the United States.* Reprinted from Hearings before the Subcommittee on Economic Statistics of the Joint Economic Committee, 85th Congress, Oct. 29 and 30, 1957. New York: National Bureau of Economic Research, Inc., 1958. $.65. Paper.

Morse, Philip M. *Queues, Inventories and Maintenance.* New York: John Wiley and Sons, Inc., 1958. $6.50.

Ohkawa, Kazushi. *The Growth Rate of the Japanese Economy.* Tokyo: Kinokuniya Book Store Co., Ltd., 1957.

Public Health, Department of, State of California. *Health in California. Summary of the California Health Survey of 1954–55.* Sacramento: California State Printing Office, Documents Section, North Seventh St. and Richards Blvd. $1.50. Paper.

Saulnier, R. J., Halcrow, Harold G., and Jacoby, Neil H. *Federal Lending: Its Growth and Impact.* Occasional Paper 58. New York: National Bureau of Economic Research, 1957. $1.00. Paper.

Smart, W. M. *Combination of Observations.* New York: Cambridge University Press, 1958. $6.50.

Souther, James W. *Technical Report Writing.* New York: John Wiley and Sons, Inc., 1957. $2.95. Paper.

Steck, G. P. *Upper Confidence Limits for the Failure Probability of Complex Networks.* Albuquerque, New Mexico: Sandia Corporation, 1957. Available from the Office of Technical Services, Dept. of Commerce, Washington 25, D. C. $1.25. Paper.

Strecker, Heinrich. *Moderne Methoden in der Agrarstatistik.* Wurzburg: Physica-Verlag, 1957. Paper.

Taeuber, Conrad, and Taeuber, Irene B. *The Changing Population of the United States.* New York: John Wiley and Sons, Inc., 1958. $7.75.

United States Dept. of Agriculture. *Agricultural Prices and Parity,* Volume 1. Agriculture Handbook No. 118. Washington, D. C.: Government Printing Office, 1957. $.50. Paper.

United States Dept. of Commerce, Bureau of the Census. *County and City Data Book, 1956.* Washington, D. C.: Government Printing Office, 1957. $4.50.

Wharton School of Finance and Commerce, University of Pennsylvania. *Study of Consumer Expenditures, Incomes and Savings.* Statistical Tables, Urban U. W., 1950. Philadelphia, Pa.: University of Pennsylvania, 1957. $6.00.

AMERICAN STATISTICAL ASSOCIATION

REPORT OF THE BOARD OF DIRECTORS, 1957

In continuation of the policy of accruing surplus whenever conditions warrant, to accumulate a total equal to one year's income, the Board reports a moderate addition for 1957 to the cumulative reserve. As is the case in many associations, rising costs may preclude further sizeable additions for some time unless the Association increases its dues as several others have done recently, or finds other sources of income.

Publications

The area of publications is causing considerable concern on the part of many interested members. The necessity of providing a forum for the many varied disciplines using statistical theory and applications will require in coming years a growing expenditure for different types of publications.

The Committee on Publications Policy—As part of its study of the whole problem of the Association's publications, the Committee prepared a letter that was sent in August over President Leonard's signature to members of the Board of Directors, the Council and officers of the five Sections, a total of approximately 60 persons. The letter requested views concerning various aspects of the problem including, for example, the matter of the number and kinds of articles in the *Journal* which are of interest to the members of the several sections. The letter also dealt with the Proceedings of the Business and Economic Statistics Section, *The American Statistician*, and the possibility of separate journals or issues of the *Journal* devoted to a specific area of interest. Following discussion of the replies, the Board and Council authorized that each Section appoint a liaison editor to the *Journal*. These have now been appointed and cooperation with them on the part of all members will help to make this important step successful. (The names of these liaison editors appear in each issue of the *Journal* on the editor's page.) The complete statement of the Board and Council's action appears in the October 1957 issue of *The American Statistician*.

Following the completion of this study by the Committee, William S. Connor asked to be relieved as Chairman due to the demands of other duties. Mr. Connor remains as a member of the Committee and Almarin Phillips, former Editor of *The American Statistician*, was appointed to succeed him. The Board has authorized a statistical study of the opinions of *Journal* users and the Committee is preparing the questionnaire that will be sent to a sample of the membership. The Association's publications are one of its most important activities. They merit careful re-examination at frequent intervals if the Association is to continue to serve the many diverse interests and fields where statistics is a tool of growing importance.

1958 Membership Directory—In the fall, a brief questionnaire was sent to a 10% sample of the members, both domestic and foreign, requesting opinions on the inclusion of various items in the next edition of the Directory. As a result of the replies received, certain new items of information about each member will be added to the alphabetical section of the book. The last edition of the Directory was published in 1954.

Proceedings of the Business and Economic Statistics Section—The publication of a 1957 volume of proceedings was authorized by the Board prior to the Annual Meeting last September. This edition which contains many papers presented in sessions of the Business and Economic Statistics Section at Atlantic City, will be available early in 1958. The 1955–56 Proceedings, which appeared early in 1957, have now been sold. There were approximately 1,200 copies printed and, consequently, almost all costs were recovered.

Grants

A grant of $9,000 to cover a three-year period was received in the spring of 1957 from the Carnegie Corporation. Its purpose is to pay travel expenses to international meetings for official delegates from the Association. A Committee was authorized by the Board to set forth the stipulations governing individual applications for travel assistance. The complete details concerning applications for the meeting of the International Statistical Institute in Brussels in 1958 are printed in the December 1957 issue of *The American Statistician*. Six other societies were each granted a like amount at the same time as ASA

and the Social Science Research Council received $150,000 for the same purpose. Appropriate liaison is being worked out with other societies, with the Council and with the National Science Foundation.

As reported previously, a grant of $19,640 was received from the Ford Foundation in 1956 for the preparation of an index to the *Journal of the American Statistical Association* to cover Volumes 35 through 50. The work is under the immediate supervision of Boyd Harshbarger at Virginia Polytechnic Institute. An Advisory Committee of the Association chaired by William G. Cochran assists Mr. Harshbarger and his associates.

Other Activities

New Twin Cities (Minn.) Chapter—A charter has been granted to a new Chapter of the Association in the Twin Cities area of Minnesota. The number of chapters is now 36.

Weiss Memorial Fund—The fund administered by the University of Pennsylvania in memory of Samuel Weiss, has reached a total of slightly over $2,500. The money will be used for student aid on the graduate level.

Reduction of Dues for Cuba and Mexico—Beginning in 1958, the dues for members in Cuba and Mexico have been reduced to $5.00 per year. This brings the payment into line with other countries.

Future Annual Meetings

The Annual Meetings set to date are as follows:

1958	Chicago, Ill.	Congress Hotel	December 27–30	Joint with Economic and Allied Associations
1959	Wash., D. C.	Shoreham Hotel	December 27–30	Joint with Economic and Allied Associations
1960	California	Palo Alto (Stanford)	Late August	
1961	New York City	Roosevelt Hotel	December 27–30	Joint with Economic and Allied Associations

The Atlantic City meeting in September 1957 was very well attended. More than 1,130 persons registered for the sessions and the hotel had to arrange for quarters nearby for many who were too late for reservations. The meeting was noted also for the largest number of individual sessions ever presented at an Annual Meeting of the Association.

The Council of the Association, enlarged in 1956, is helping measurably to improve member participation in the Association's activities. During the year, officers traveling in various parts of the country participated in Chapter meetings and plans for greater activity along these lines are being made for the forthcoming year.

REPORT OF THE SECRETARY-TREASURER, 1957

The surplus for 1957 amounts to $7,700. Added to the cumulative surplus to date, the Association starts 1958 with approximately $45,500 in reserve funds. The 1957 surplus results more from a reduction in expenditures than increases in income. This is particularly true in the printing budget of the *Journal*, though some other expense items are smaller as well. Some income items show an increase over budget, the most significant being advertising and *Journal* sales.

Specifically, the 1957 budget was set at $71,925.00 for income and $71,845.00 for expense. Actual income is about $75,400, while expense for the year totals approximately $67,600. Note that the membership directory actual expense allocated for 1957 has been raised to $6,000 instead of $3,000 as budgeted. This is the expense for the 1958 directory on which preparation has recently begun. (Total Directory cost is expected to be $9,000.) Although it has been the usual policy to spread directory expense over a three-year period, a larger share has been allocated to 1957 instead of placing this $3,000 into surplus. Therefore, the final expense for the directory will be taken in the 1958 budget leaving 1959 free for other considerations. (The Journal Index to Volumes 35–50 may be ready for printing in 1959.)

For a number of years, it has been reported in the annual audit that dues income exceeded the amount budgeted. The situation for 1957 is reversed—actual dues income fell short of budget. It is significant that for the first time in ten years the number of

members actually decreased. There are several possible explanations for this: the fact that ASA has a larger "fringe" membership than other societies, with these marginal-interest members dropping out for economic or other reasons; the lack of sustained membership promotion drives in 1957 (a condition that has already been corrected for 1958); the problem concerning publications which has been pointed up in the report of the Board of Directors.

The actual count of members at the beginning of 1957 was 5,691. During 1957, 530 new members joined the Association and 32 reinstated their membership after a lapse of one or more years. However, 136 persons resigned or died and 450 were dropped for nonpayment of dues. This last figure is considerably larger than preceding years when it has generally approximated 300. Thus, the Association starts 1958 with 5,667 members.

As has been mentioned above, promotion campaigns in 1958 will be actively pursued. The results of the first invitational mailings have already been felt—the number of new members joining during approximately the last month of 1957 and credited to 1958 is more than twice the number joining in the comparable period a year ago. (Subsequently, the upward membership trend has again accelerated and membership now totals over 6,000.) *Journal* subscriptions (to libraries, business firms, etc.) continued to increase. The number for this year is 1,784 as compared to 1,657 in 1956. Campaigns for new *Journal* subscribers (as distinct from membership) have been started as well.

Income for 1958 has been budgeted at 74,970, with expenses at the same figure. In line with previous policy discussions in the Council and Board, a sum of $2,620 is being set aside from reserve to meet contingent *Journal* expense. This leaves the *Journal* with the budgetary latitude that is felt to be desirable. When additional suitable articles are available, the funds for an increased number of pages would be available also. Should the *Journal* printing expense as well as other expense items be substantially under the 1958 budget, it is possible that we shall end the forthcoming year with another small addition to surplus.

One other point should be noted concerning the 1958 expense budget. Due to the addition of another part-time person to the national office staff and, to a lesser extent, raises in salary, the expense item for salaries has increased. The growth in size and activities of ASA over the past five years alone has necessitated hiring extra help. The full-time and part-time staff thus is once more at about the same size as that in 1950 when membership numbered approximately 4,200. Salaries paid are in keeping with the wages paid for similar services elsewhere in Washington.

Every effort will be made in 1958 to increase income while practicing stringent economies on expenses. The reserve fund must be increased in anticipation of future contingencies.

December, 1957

DONALD C. RILEY
Secretary-Treasurer

JAMES G. JESTER
CERTIFIED PUBLIC ACCOUNTANT
TOWER BUILDING
WASHINGTON 5, D. C.

April 22, 1958

To the Board of Directors of the
American Statistical Association:

I have examined the financial statements of the American Statistical Association relating to the year ended December 31, 1957. My examination was made in accordance with generally accepted auditing standards and, accordingly, included such tests of the accounting records and other auditing procedures as were considered necessary in the circumstances, except that I did not check the membership and subscription records in detail, or make any independent verification of the inventory of back numbers of the *Journals* the office records of which are based, in part, on data assembled in prior years.

In my opinion, the accompanying statements present fairly the financial position of the American Statistical Association on December 31, 1957, and the results of its operations for the year then ended, in accordance with generally accepted accounting principles applied on a basis consistent with that of the preceding year.

JAMES G. JESTER

Exhibit A

AMERICAN STATISTICAL ASSOCIATION

COMPARATIVE BALANCE SHEET

Assets

	December 31,	
	1957	*1956*
Cash in banks and on hand	$ 30,064.70	$20,133.63
Accounts receivable	2,438.81	912.84
Investments:		
United States Savings bond, Series G, due 1962, at cost	3,100.00	3,100.00
United States Treasury 2½% bonds, $7,000 face amount, due November 1961, at cost	9,566.88	6,700.94
Federal Home Loan Bank Board	5,000.00	
Federal National Mortgage Association	5,000.00	
Funds on deposit with savings and loan institutions—Schedule I	54,392.00	56,909.09
Inventories:		
Old *Journals*, at approximate cost	3,113.49	2,922.00
Monograph on Kinsey Report, at cost	2,746.61	2,848.87
Emblems, at cost	136.50	217.50
Furniture and fixtures, at cost, less accumulated depreciation	2,234.08	1,660.42
Deferred charges:		
Postage and supplies	400.33	527.35
Annual meeting expense	189.61	382.00
Prepaid rent		200.00
Prepaid insurance and fidelity bond	604.00	202.02
Pension Plan premium	1,181.32	
Total Assets	$120,168.33	$96,716.66

Liabilities and Net Worth

Accounts payable	$ 11,304.44	$ 5,172.65
Deferred income (collections applicable to subsequent years):		
Dues	$ 26,339.00	$20,740.00
Subscriptions	7,792.40	7,078.40
Index to the *Journal* grant	13,037.10	$19,640.00
Carnegie Institution Travel Fund Grant	8,931.72	
Other	466.84	2,467.33
	$ 56,567.06	$49,925.73
Deferred credit:		
Appropriation in 1957 toward publication costs of the 1958 Membership Directory	$ 3,000.00	
Net Worth:		
Life Membership reserve	$ 3,743.95	$ 3,843.47
Surplus—Exhibit B	45,552.88	37,774.81
Total net worth	$ 49,296.83	$41,618.28
Total Liabilities and Net Worth	$120,168.33	$96,716.66

Exhibit B

AMERICAN STATISTICAL ASSOCIATION

COMPARATIVE STATEMENT OF INCOME AND SURPLUS

	Year ended December 31,	
	1957	*1956*
Income:		
Dues—Old members, including Life members	$41,523.40	$40,432.26
—New members	3,873.75	4,683.00
Subscriptions—*Journal*	13,336.73	12,441.53
—*American Statistician*	541.42	504.45
Advertising—*Journal*	2,903.55	1,661.08
—*American Statistician*	918.37	699.76
Sales—*Journal*	2,531.87	1,308.82
—*American Statistician*	197.97	176.45
—Kinsey Report Monograph	231.66	602.00
—Business and Economic Section Proceedings	3,018.70	456.69
—Emblems, less cost of sales	47.85	62.67
—Other	127.25	57.58
Mailing list income	1,418.45	1,473.91
Interest income	2,593.68	1,631.69
Annual meeting, including Exhibits	2,137.88	2,231.84
Miscellaneous	19.86	99.58
Total income	$75,422.39	$68,523.31
Expenses:		
Salaries	$17,578.43	$15,418.62
Pension Plan expense	393.78	
Publications—Schedule I	33,973.42	26,991.40
Promotion	1,650.53	45.00
Rent	2,400.00	2,400.00
Travel and Secretary's expense	1,142.43	732.55
Supplies and office expenses	2,724.69	2,414.66
Postage	1,973.03	2,071.07
Telephone and telegraph	810.86	887.45
Accounting services	970.00	970.00
Committee expense	612.51	741.91
Annual meeting expense	1,856.98	1,095.66
Miscellaneous expenses—Schedule I	1,557.66	1,790.00
	$67,644.32	$55,558.32
Excess of income over expense for the year	$ 7,778.07	$12,964.99
Add: Surplus account at beginning of year	37,774.81	24,809.82
Surplus account at end of year—Exhibit A	$45,552.88	$37,774.81

Schedule I

AMERICAN STATISTICAL ASSOCIATION

	Year ended December 31, 1957	1956
Publications:		
Journal—Printing	$16,191.05	$15,078.02
—Abstracts	750.00	750.00
—Editorial expense	2,298.67	2,493.26
	$19,239.72	$18,321.28
American Statistician	8,382.60	6,652.67
Monographs—Kinsey report	102.26	190.17
Business and Economic Section Proceedings	3,150.24	40.18
Membership Directory (includes $3,000 budgeted in 1957 toward publication costs of the 1958 Membership Directory)	3,098.60	1,787.10
Total publications—Exhibit B	$33,973.42	$26,991.40
Miscellaneous Expenses:		
Depreciation	$ 371.08	$ 308.71
Dues to other organizations	100.00	123.25
Taxes	477.86	413.38
Repairs and maintenance	169.35	115.71
Insurance	253.97	251.91
Other	185.40	577.04
Total miscellaneous expenses—Exhibit B	$ 1,557.66	$ 1,790.00
Funds on deposit with savings and loan institutions:		
American Savings and Loan Association	$ 5,580.84	$ 5,390.53
City Federal Savings & Loan Association	10,000.00	10,000.00
Fidelity Federal Savings & Loan Association	9,482.44	9,113.33
Guardian Savings & Loan Association	5,000.00	5,015.63
Hyattsville Building Association	1,132.26	1,092.57
Interstate Building Association		8,140.48
Liberty Savings & Loan Association	3,714.96	3,605.70
Marshall Savings & Loan Association	5,000.00	
National Permanent Building Association		5,437.53
Piedmont Federal Savings & Loan Association	9,481.50	9,113.32
Trans-Bay Federal Savings & Loan Association	5,000.00	
Total funds on deposit with savings and loan institutions	$54,392.00	$56,909.09

IBM STATISTICIANS

Excellent opportunity for statisticians with background in engineering field. M.S. in Statistics, with major work in math statistics. Minimum of two years' experience in statistical work, preferably with engineering applications. Thorough knowledge of statistical distribution functions. To handle anlysis-of-variance and multiple-regression type problems; to design experiments for various engineering applications; to select the form of statistical analysis of greatest value; to give statistical support to engineering departments in such areas as flight test, reliability analysis, and human factors engineering; to develop statistical programs for the IBM 704.

Advantages of IBM

A recognized leader in the electronic computer field . . . stable balance of military and commercial work . . . advancement on merit . . . company-paid relocation expenses . . . liberal company benefits . . . salary commensurate with ability and experience.

Immediate openings at Owego and Endicott, N.Y.

WRITE, outlining qualifications and experience, to:

Mr. R. E. Rodgers, Dept. 577R
Technical Recruitment Mgr.
IBM Corp., 590 Madison Avenue
New York 22, New York

DATA PROCESSING
ELECTRIC TYPEWRITERS
MILITARY PRODUCTS
SPECIAL ENG'G. PRODUCTS
SUPPLIES
TIME EQUIPMENT

IBM INTERNATIONAL BUSINESS MACHINES CORPORATION

Ph.D. MATHEMATICIANS

Our Systems Analysis Department has openings for Ph.D. Mathematicians who have a lasting interest in some fields of Mathematics such as Algebra, Geometry, Measure Theory, Topology, Differential Equations, Numerical Analysis, Statistics, or Probability Theory. Association will be with others with similar educational background engaged in the study of challenging problems vital to the nation's welfare.

The work is of an independent research nature with supervision at a minimum and by people of similar interests who are aware of the needs of research personnel.

Sandia Corporation, engaged in research and development of nuclear weapons, is located in Albuquerque, N.M.—a metropolitan city of 208,000 famous for its splendid climate and its outstanding recreational and cultural attractions. Excellent working conditions and liberal employee benefits, including generous vacations and holidays; retirement, hospital, and insurance plans.

Please send résumé to Staff Employment Division 501

SANDIA CORPORATION
ALBUQUERQUE, NEW MEXICO

BIOMETRIKA

Vol. 45, Parts 1 & 2 June, 1958

CONTENTS

Memoirs:

ARMITAGE, P. Numerical studies in the sequential estimation of a binomial parameter.

LESLIE, P. H. A stochastic model for studying the properties of certain biological systems by numerical methods.

SKELLAM, J. G. On the derivation and applicability of Neyman's Type A distribution.

BLISS, C. I. & OWEN, A. R. G. Negative binomial distributions with a common *k*.

BRASS, W. Simplified methods of fitting the truncated negative binomial distribution.

COX, D. R. The interpretation of the effects of non-additivity in the latin square.

ASHFORD, J. R. Quantal responses to mixtures of poisons under conditions of simple similar action—the analysis of uncontrolled data.

MOORE, P. G. Some properties of runs in quality control procedures.

WILLIAMS, E. J. Simultaneous regression equations in experimentation.

RUSSELL, T. S. & BRADLEY, R. A. One-way variances in a two-way classification.

PLACKETT, R. L. Studies in the history of probability and statistics. VII. The principle of the arithmetic mean.

BROWN, R. L. & FEREDAY, F. Multivariate linear structural relations.

OLKIN, I. Multivariate ratio estimation for finite populations.

BARTON, D. E., DAVID, F. N. & MALLOWS, C. L. Non-randomness in a sequence of two alternatives. I. Wilcoxon's and allied test statistics.

GOODMAN, L. A. Simplified runs tests and likelihood ratio tests for Markoff chains.

LEIPNIK, R. B. Moment generating functions of quadratic forms in serially correlated normal variables.

SAW, J. G. Moments of sample moments of censored samples from a normal population.

HERDAN, G. The relation between the dictionary distribution and the occurrence distribution of word length and its importance for the study of quantitative linguistics.

MACKENZIE, J. K. Second paper on statistics associated with the random disorientation of cubes.

WAUGH, W.A.O'N. Conditioned Markov processes.

Miscellanea: Contributions by D. E. BARTON, D. E. BARTON & F. N. DAVID, H. BERGSTROM, C. W. CLUNIES-ROSS, H. R. B. HACK, M. HALPERIN & S. W. GREENHOUSE, G. HERDAN, A. HUITSON, N. L. JOHNSON, J. E. KERRICH, R. D. LORD, R. MAURICE.

Corrigenda: F. N. DAVID & M. G. KENDALL: A. J. MAYNE: E. L. CROW.

Reviews

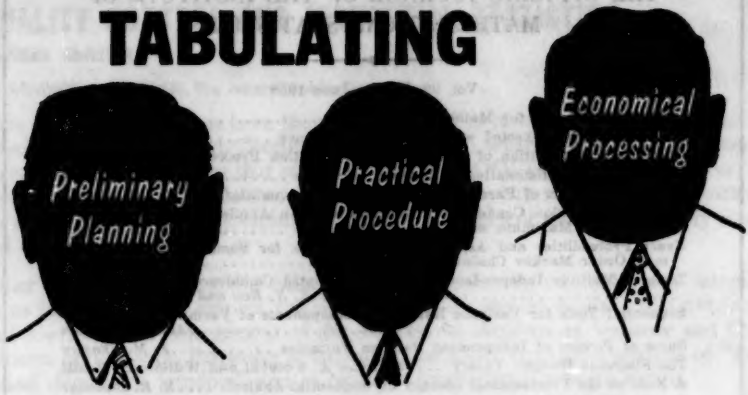
Other Books Received

The subscription, payable in advance, is 54/- (or $8.00) per volume (Including postage). Cheques should be made payable to *Biometrika*, crossed *"a/c Biometrika Trust"* and send to the Secretary, Biometrika Office, Department of Statistics, University College, London, WC1. All foreign cheques must be drawn on a Bank having a London agency.

Issued by the Biometrika Office, University College, London

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

The IDEA Factors Behind MARKET RESEARCH TABULATING



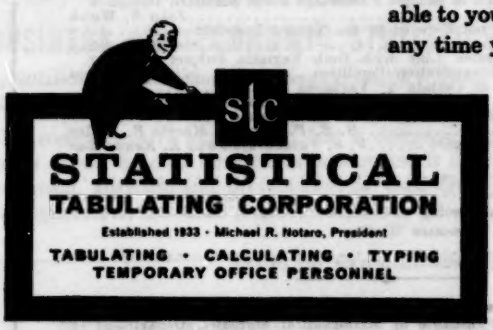
STATISTICAL service begins long before a button is pushed.

You get preliminary assistance in resolving your ideas . . . in translating sound thinking into well-planned questionnaires for the most practical and economical processing.

There is always a best way to handle any assignment and STATISTICAL can help you apply it through long experience in methods and procedures.

The same careful approach is used in processing data to assure highest quality in market information. Strict controls are maintained every step of the way from editing and coding to finished report.

And this professional service is available to you days, nights, week-ends—any time you need it.



Just call our nearest office

CHICAGO, *53 W. Jackson, HArrison 7-4500* • **NEW YORK,** *80 Broad St., WHitehall 3-8383*
ST. LOUIS, *411 N. Tenth St., MAin 1-7777* • **NEWARK,** *National Newark Bldg., MArket 3-7636*
CLEVELAND, *1375 Euclid Ave., SUperior 1-8101* • **LOS ANGELES,** *541 S. Spring St., MAd. 6-8201*
KANSAS CITY, *20 W. 9th St., GRand 1-3629*

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

The Annals of Mathematical Statistics

THE OFFICIAL JOURNAL OF THE INSTITUTE OF MATHEMATICAL STATISTICS

Vol. 29, No. 2—June 1958

Information Theory for Mathematicians ..... *J. Wolfowitz*
Some Problems Connected with Statistical Inference ..... *D. R. Cox*
Asymptotic Distribution of Stochastic Approximation Procedures *Jerome Sacks*
Admissibility for Estimation with Quadratic Loss ..... *Samuel Karlin*
On the Estimation of Parameters Restricted by Inequalities ..... *H. D. Brunk*
Intersection Region Confidence Procedures with an Application to the Location of the Maximum in Quadratic Regression ..... *David L. Wallace*
Exact Probabilities and Asymptotic Relationships for Some Statistics From m-th Order Markov Chains ..... *Leo A. Goodman*
Tests of Multiple Independence and the Associated Confidence Bounds .... *S. N. Roy and R. E. Bargmann*
Sequential Tests for Variance Ratios and Components of Variance .... *Allan Birnbaum*
Sums of Powers of Independent Random Variables ..... *J. M. Shapiro*
The Staircase Design: Theory .... *Franklin A. Graybill and William E. Pruitt*
A Note on the Fundamental Identity on Sequential Analysis .... *R. R. Bahadur*
On the Statistical Treatment of Stochastic Processes .... *Isidore Fleischer and Anthony Kooharian*
Some Remarks on Sampling with Replacement *Dees Raj and Salem H. Khamis*
A Distribution-free Upper Confidence Bound for $Pr\{Y < X\}$, Based on Independent Samples of X and Y ..... *Z. W. Birnbaum and R. C. McCarty*
On the Integrodifferential Equation of Takacs. I ..... *Edgar Reich*
Probabilities of Hypotheses and Information-Statistics in Sampling from Exponential-class Populations ..... *Morton Kupperman*

Notes:

On the Distribution of 2 x 2 Random Normal Determinants *W. L. Nicholson*
A Smooth Invertibility Theorem ..... *J. W. Tukey*
On the Inversion of the Sample Covariance Matrix in a Stationary Autoregressive Process ..... *M. M. Siddiqui*
A Problem of Berkson, and Minimum Variance Orderly Estimators .... *J. W. Tukey*
An Elementary Theorem Concerning Stationary Ergodic Processes .... *Leo Breiman*
A Test of Fit for Multivariate Distributions ..... *Lionel Weiss*
Tables for Obtaining Non-parametric Tolerance Limits ..... *Paul N. Somerville*
Nonparametric Estimation of Sample Percentage Point Standard Deviation ..... *John E. Walsh*
A Uniqueness Property not Enjoyed by the Normal Distribution .... *George P. Steck*
Estimation of a Regression Line with Both Variable Subject to Error Under an Unusual Identification Condition ..... *Herman Rubin*
On the Decomposition of Certain χ^2 Variables .... *Robert V. Hogg and Allen T. Craig*
A Note on the Generation of Random Normal Deviates .... *G. E. P. Box and Mervin E. Muller*
On a Theorem in Metric Spaces ..... *V. S. Varadarajan and R. Ranga Rao*

Abstracts of Papers .....
News and Notices .....
Report of the Ames, Iowa Meeting .....
Report of the Gatlinburg, Tennessee Meeting .....
Publications Received .....

Address orders for subscriptions and back numbers to Professor George E. Nicholson, Jr., Secretary, Institute of Mathematical Statistics, Department of Statistics, University of North Carolina, Chapel Hill, North Carolina.

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

Popular Irwin Texts

QUALITY CONTROL AND INDUSTRIAL STATISTICS

REVISED EDITION

By ACHESON J. DUNCAN, The Johns Hopkins University

This outstanding text has been thoroughly revised to bring it in line with all recent developments in the field. Specific changes include: separation of the discussion of acceptance sampling from that of rectifying (i.e. AOQL) inspection, addition of a special chapter on multiple sampling plans, and complete rewriting of the material on acceptance sampling by variables. The new government standard, Military Standard 414, is described.

Other improvements include enlargement of the section on continuous sampling plans into a complete chapter, the addition of a chapter on Analysis of Covariance, and rewriting and enlargement of the chapters on Analysis of Variance and Components of Variance. Particular attention has been devoted to the analysis of "individual degrees of freedom" and to multiple comparisons between means.

Some 400 questions and problems are provided, many of them new in the Revised Edition. A complete Solutions Manual will be available to adopters.

WORKBOOK IN BUSINESS STATISTICS

FOURTH EDITION

By LOUIS F. HAMPEL, United Air Lines

This popular workbook is designed for classroom and laboratory use in the study of business statistics. The new Fourth Edition has been thoroughly revised and brought up to date. Emphasis is on short problems that do not require the use of calculating machines. This Workbook may be used with any standard text in the field. A complete Solutions Manual is available to adopters.

BUSINESS AND ECONOMIC STATISTICS

By WILLIAM A. SPURR, Stanford University, LESTER S. KELLOGG, Deere and Company, and JOHN H. SMITH, The American University

This comprehensive introductory text is designed for the student who must use statistics in any field of business or economics. It deals in elementary terms with the collection, analysis, and presentation of data. Short-cut graphic methods are used throughout and statistical jargon is avoided.

Write for Examination Copies

RICHARD D. IRWIN, INC. • HOMEWOOD, ILLINOIS

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

BIOMETRICS

Journal of the Biometric Society

Vol. 14, No. 2 CONTENTS June 1958

A Generalized Class of Contagious Distributions (John Gurland); On the Probability of Survival of Bacteria in Sea Water (Eugene K. Harris); Efficient Estimation of the Relationship Between Plot Size and the Variability of Crop Yields (W. H. Hatheway and E. J. Williams); Maximum Likelihood Estimation from Incomplete Data (H. O. Hartley); Estimation of Relative Frequencies of Four Sperm Types in Drosophila Melanogaster (Marvin A. Kastenbaum); The Effect of Sterilised Males on a Natural Tsetse Fly Population (H. R. Simpson); Estimation of Missing Values for the Analysis of Incomplete Data (G. N. Wilkinson); Recent Experiences with Area Sampling for Agricultural Statistics (R. E. Vickery); Query: Orthogonal Coefficients for Unequal Intervals (A. Grandage).

Biometrics is published quartely. Its objects are to describe and exemplify the use of mathematical and statistical methods in biological and related sciences, in a form assimilable by experimenters. Members of the American Statistical Association may subscribe through the Association at the rate of $4 yearly. The annual non-member subscription rate is $7. Inquiries, orders for back issues and non-member subscriptions should be addressed to:

BIOMETRICS
Department of Statistics
Virginia Polytechnic Institute
Blacksburg, Virginia, U. S. A.

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

New McGRAW-HILL Books

COMPUTABILITY AND UNSOLVABILITY

By MARTIN DAVIS, Rensselaer Polytechnic Institute. *McGraw-Hill Series in Information Processing and Computers.* 214 pages, $7.50

Concerned with the existence of algorithms or effective computation procedures for solving various problems. It deals with the general theory of computability and the application of the subject to algebra, number theory, and symbolic logic. Here is the first connected presentation of this theory offered from the point of view of Turing machines.

HIGH-SPEED DATA PROCESSING

By C. C. GOTLIEB and J. N. P. HUME, both of the University of Toronto. *McGraw-Hill Series in Information Processing and Computers.* 338 pages, $9.50

A basic, comprehensive treatment of the important principles and general techniques of processing data at high speeds. It shows how data processors work, how to use them, and what their advantages are. Coding and programming methods are included and examples of typical applications of high-speed data processing are shown.

APPLIED STATISTICS FOR ENGINEERS

By WILLIAM VOLK, Hydrocarbon Research, Inc., Princeton, N.J. *McGraw-Hill Series in Chemical Engineering.* 354 pages, $9.50

Provides enough background and examples of statistical problems to enable practicing engineers to apply statistical analysis to their data. Emphasizing applications and providing many illustrative examples, it deals with the treatment of engineering data for correlation, precision, and analysis of experimental factors. Each chapter is complete in itself.

LINEAR PROGRAMMING: Methods and Applications

By SAUL I. GASS, U.S.D.A. Graduate School, Washington, D.C. Ready in Fall

This excellent introductory text for statistical courses presents the theoretical, computational, and applied areas of linear programming—explaining the problems it can solve and the implementing of techniques to formulate and solve them. It covers the revised simplex method, parametric linear programming, and a summary of digital computer codes.

Send for copies on approval

McGRAW-HILL BOOK COMPANY, INC.

330 WEST 42nd STREET, NEW YORK 36, N.Y.

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

ESTADISTICA

Journal of the
Inter American Statistical Institute

Vol. XV, No. 58 **CONTENTS** **March 1958**

ARTICLES

Addresses of the V Session of the Committee on Improvement of National Statistics (COINS)—Discursos de la V Sesión de la Comisión de Mejoramiento de las Estradísticas Nacionales (COINS) .....W. P. Bermúdez and R. D. Narain

Un Informe sobre Perspectivas (traducción) .....William R. Leonard

Comentarios sobre Geografía y Cartografía para Propósitos Censales, con Enfasis en la Preparación de Mapas para los Censos de América de 1960 (traducción) .....Clarence E. Batschelet y Alford Archer

Contributions a la Determination du Niveau de la Mortalité au Brésil .....Ernani Thimoteo de Barros

Objective Methods for Forecasting Florida Citrus Production .....Bruce W. Kelly

Internal Migration in Mexico .....Nathan L. Whetten and Robert G. Burnight

Las Contribuciones de Karl Pearson (traducción) .....Helen M. Walker

Special Features. Legal Provisions. International Resolutions Relating to Statistics. Institute Affairs. Statistical News. Publications.

Published quarterly **Annual subscription price $3.00 (U.S.)**

Inter American Statistical Institute
Pan American Union, Washington 6, D.C.

ECONOMETRICA

Journal of the Econometric Society

Vol. 26, No. 2, April, 1958

CONTENTS

R. DUNCAN LUCE: A Probabilistic Theory of Utility

ROBERT J. WOLFSON: An Econometric Investigation of Regional Differentials in American Agricultural Wages

WILLIAM J. BAUMOL: Topology of Second Order Linear Difference Equations with Constant Coefficients

JEREMY J. STONE: An Experiment in Bargaining Games

K. J. ARROW AND M. NERLOVE: A Note on Expectations and Stability

NANDA K. CHOUDHRY: An Econometric Analysis of the Import Function for Burlap in the U. S. A.

H. O. CARTER AND H. O. HARTLEY: A Variance Formula for Marginal Productivity Estimates Using the Cobb-Douglas Function

J. VAN IJZERN: A Note on the Useful Properties of Mr. Stuvel's Index Numbers

REPORT OF THE ATLANTIC CITY MEETING

BOOK REVIEWS

ANNOUNCEMENTS AND NOTES

THE ECONOMETRIC SOCIETY
YALE UNIVERSITY
NEW HAVEN, CONN.

AN INVITATION TO JOIN ORO

Pioneer In Operations Research

Operations Research is a young science, earning recognition rapidly as a significant aid to decision-making. It employs the services of mathematicians, physicists, economists, engineers, political scientists, psychologists, and others working on teams to synthesize all phases of a problem.

At ORO, a civilian and non-governmental organization, you will become one of a team assigned to vital military problems in the area of tactics, strategy, logistics, weapons systems analysis and communications.

No other Operations Research organization has the broad experience of ORO. Founded in 1948 by Dr. Ellis A. Johnson, pioneer of U. S. Opsearch, ORO's research findings have influenced decision-making on the highest military levels.

ORO's professional atmosphere encourages those with initiative and imagination to broaden their scientific capabilities.

ORO starting salaries are competitive with those of industry and other private research organizations. Promotions are based solely on merit. The "fringe" benefits offered are ahead of those given by many companies.

The cultural and historical features which attract visitors to Washington, D. C. are but a short drive from the pleasant Bethesda suburb in which ORO is located. Attractive homes and apartments are within walking distance and readily available in all price ranges. Schools are excellent.

For further information write:
Professional Appointments

OPERATIONS RESEARCH OFFICE

ORO The Johns Hopkins University

6935 ARLINGTON ROAD
BETHESDA 14, MARYLAND

THE REVIEW OF ECONOMICS AND STATISTICS

Pubished by the Department of Economics, Harvard University

Editor: Seymour E. Harris*

Associate Editors: Abram Bergson, Edward H. Chamberlin, Robert Dorfman, James S. Duesenberry, John T. Dunlop, John Kenneth Galbraith,* Alexander Gerschenkron, Gottfried Haberler,* Alvin H. Hansen,* Carl Kaysen, Wassily Leontief,* John Lintner,* Edward S. Mason, Sumner H. Slichter, Arthur Smithies, John H. Williams.

* Members of Managing Board.

VOLUME XL MAY 1958 NUMBER 2

The State of Economic Science ..... *Wassily Leontief*
Forces Widening Occupational Wage Differentials *Richard Perlman*
The Theory of the Responsiveness of Hours of Work to Changes in Wage Rates ..... *Franklee Gilbert* and *Ralph W. Pfouts*
The Renaissance of Soviet Statistics ..... *Harry Schwartz*
Communist China's National Product in 1952 ... *Alexander Eckstein* with the assistance of *Y. C. Yin* and *Helen Yin*
The Operation of Trade within the Soviet Bloc ..... *Nicolas Spulber* and *Franz Gehrels*
Job Analysis and Historical Productivities in the American Cotton Textile Industry: A Study in Methodology ..... *T. Y. Shen*
The Stability of Business Capital Outlays ..... *Robert Lindsay*

NOTES AND BOOK REVIEWS

The Review of Economics and Statistics *is published quarterly.*

Annual subscription price: $6.00.

Order from the Harvard University Press, Cambridge 38, Mass.

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

Statistical Yearbook 1957

A comprehensive collection of international statistics relating to economic, demographic, financial, and social subjects. There are 191 tables, most of which cover a twenty-year period ending in 1956 or mid-1957. This volume presents an important new chapter on international economic aid, total assistance related to population and per capita gross national product, and the contributions of governments to international technical assistance and relief agencies. A feature of this issue is the large number of statistics it contains for the USSR and East European countries. Other new tables relate to ocean freight rates, discount rates of central banks, and number of translations of books by subject groups and original languages. *A United Nations publication.*

Sales No. 1957. XVII.1: cloth $8.00, paper $6.50

Demographic Yearbook 1957

This issue features mortality statistics in a historical series that analyzes and links the rates of death to age, sex, and cause, back through the 1930s. A compendium of official demographic statistics, it contains information on time trends of population estimates, births, deaths, stillbirths, marriages, and divorces. Contributions are from over 200 sovereign countries, non-self-governing territories, and dependencies throughout the world. New features include a series for city population and percentage rate of population growth from 1953 to 1956. An analytical chapter on the decline in mortality throughout the world will supplement the regularly published "Technical Notes" that describe the limitations in comparability of the statistics. *A United Nations publication.*

Sales No. 1957. XIII.1: cloth $8.00, paper $6.50

Yearbook of International Trade Statistics 1956

Volume I discusses total imports and exports (in national currency and U.S. dollars), trade in gold, indices of quantum and unit value, currency conversion factors, trade in principal commodities (where possible according to the Standard International Trade Classification), and trade by principal countries of provenance and destination. For each of the 115 countries regional summaries with grand totals survey value of imports and exports (in U.S. dollars), provenance and destination of world exports (in matrix form), indices of quantum, and unit value for exports. Volume II contains information on provenance and destination of trade in each of the following commodity classes (values in U.S. dollars): food, beverages, tobacco, raw materials, fats and oils, fuels, chemicals, machinery, transport equipment, and other manufactured goods. *A United Nations publication.*

Sales No. 1957. XVII.6 Vol. I, Vol. II: Vol. I $7.00, Vol. II $1.50

Please order these books by sales number.

International Documents Service

COLUMBIA UNIVERSITY PRESS
2960 BROADWAY NEW YORK 27, N.Y.

THE JOURNAL OF FINANCE

Published by THE AMERICAN FINANCE ASSOCIATION

Volume XIII, No. 2 (May, 1958) includes:

Current Questions in Consumer Credit ..... *Herman E. Krooss and Peter G. Fousek*
Developments in the Mortgage Market ..... *James J. O'Leary and Mortimer Kaplan*
Effects of Taxation ... *Harold M. Somers, Walter W. Brudno, and Daniel M. Holland*
Pension Funds and the Capital Market ..... *Paul L. Howell and William C. Freund*
The International Monetary Situation ..... *Per Jacobsson*
The Money Market Outlook .... *Orvon H. Hart, Wesley Lindow, and Roy L. Reierson*

Membership dues, including $3.00 allocated to subscription in *The Journal of Finance*, are $5.00 annually. Libraries may subscribe to *The Journal* at $5.00 annually and single copies may be purchased for $1.25. Applications for membership in the American Finance Association and subscriptions to *The Journal of Finance* should be addressed to the Secretary-Treasurer, George E. Hassett, Graduate School of Business Administration, New York University, 90 Trinity Place, New York 6, N. Y.

Communication relating to the contents of *The Journal of Finance* should be addressed to the Acting Editor, Joel Segall, School of Business, University of Chicago, Chicago 37, Illinois, or to the Associate Editor, Carl A. Dauten, School of Business and Public Administration, Washington University, St. Louis 5, Missouri.

JOURNAL OF FARM ECONOMICS

Published by

THE AMERICAN FARM ECONOMIC ASSOCIATION

Editor: ROBERT L. CLODIUS
University of Wisconsin, Madison, Wisconsin

Volume XL — **May 1958** — **Number 2**

Proceedings Papers—Joint Winter Meetings

Articles

Structural Changes in the Meat Wholesaling Industry ..... *Willard F. Williams*
The Family Farm and the Land Speculator: Reflections on a Much Worked Theme ..... *E. Louise Peffer*
Some Current Developments in Futures Trading ..... *Roger W. Gray*
The Urbanization of the Open Country ..... *Kenneth C. Nobe*
Educational Requirements for Extension Workers with Suggestions on In-Service Training ..... *Harold C. Love*
The Agricultural Exemption in Interstate Trucking ..... *Celia Sperling*
Output in Relation to Input for the Agricultural Industry ..... *Earl O. Heady*
Planning Efficient Operations for Cherry Processing Plants .. *Vernon L. Sorenson*
Economies of Size in the Cattle-feeding Industry of California .... *John A. Hopkin*
Membership Attitudes of the American Farm Economic Association ..... *Harry C. Trelogan*

This journal contains additional articles, notes, book reviews, and announcement of new bulletins in agricultural economics and is published in February, May, August, November and December. Yearly subscription is $7.

Secretary-Treasurer: LOWELL S. HARDIN
Department of Agricultural Economics
Purdue University, Lafayette, Indiana

ELEMENTARY STATISTICAL METHODS as Applied to Business and Economic Data, Revised Edition

by WILLIAM A. NEISWANGER, University of Illinois

from *The Journal of the American Statistical Association:*

"Well written and widely adopted, the Neiswanger text now appears in an improved and modernized edition. This fine text has been generally known and highly regarded since its publication in 1943; the revision is both a comprehensive and meticulous job of bringing the volume up to date and perfecting the presentation. . . . The new edition is a thoughtful and careful revision in which the style, content, and emphasis of the most readable text in the field have been improved and brought up to date. . . . This review is extremely favorable and in a sense the major criticism is also a compliment. In many ways the book is too good and too complete! It makes vigorous competition for the teacher; any teacher compares favorably with a poor text but only the best will outshine Neiswanger.

"In brief it is a pleasure to report that this revision embodies many improvements over an originally very good text."

Lloyd Saville, *Duke University*

1956 *749 pages* *$7.25*

PROBABILITY, STATISTICS AND TRUTH

by RICHARD VON MISES

Second Revised English Edition by HILDA GEIRINGER

A translation of one of the internationally outstanding works on statistics, this is an analysis of the quantitative concept of probability in probability theory.

1957 *244 pages* *$5.00*

The Macmillan Company

60 FIFTH AVENUE, NEW YORK 11, N. Y.

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

STUDIES IN
The Mathematical Theory of Inventory and Production

Kenneth J. Arrow, Samuel Karlin, and Herbert Scarf

This book is the first to explore and formulate the general mathematical concepts behind efficient inventory and production control. In a series of interrelated papers, inventory polices are studied under a wide variety of assumptions likely to be met in practice. The mathematics used represent new developments in the calculus of variations and in the theory of stochastic processes. *Stanford Mathematical Studies in the Social Sciences. I.* 350 pages. 6⅞ x 10 inches. $8.75

Stanford University Press, Stanford, California

GRIFFIN'S

Statistical Monographs and Courses

Edited by M. G. Kendall

NOW AVAILABLE:

No. 1
M. H. QUENOUILLE The Analysis of Multiple Time Series, *105 pages, $4.75*

No. 2
M. G. KENDALL A Course in Multivariate Analysis, *182 pages, $4.50*

AVAILABLE SOON:

No. 3
M. H. QUENOUILLE The Fundamentals of Statistical Reasoning. *About 140 pages*

IN PREPARATION:

No. 4
A. STUART A First Course in Sampling

STECHERT-HAFNER, INC.

FOUNDED IN NEW YORK 1872

The World's Leading International Booksellers

31 EAST 10th STREET, NEW YORK 3, N. Y.

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

NEW SECOND EDITION! **NEW SECOND EDITION!**

MONEY, MEN AND MACHINES

By Waddill Catchings and Charles F. Roos

QUESTIONS FEDERAL MONETARY POLICIES	The internationally known authors of the new second edition of MONEY, MEN AND MACHINES take a firm stand against the erratic course of the Federal Reserve Board.
CRITICISM DOCUMENTED	The authors' criticism is heavily documented and they advocate a sound monetary policy to provide for a stabilized, growing economy.
DID CONGRESS INTEND IT THAT WAY?	The book discusses the serious problem whether Congress had really intended to bestow upon the FRB the economic powers which it now has.
A LOOK INTO THE FUTURE	In this new edition, three important chapters were added. Vital economic issue of this epoch and of years to come are brought to the fore.

261 pages, price $3.50

Published by

THE ECONOMETRIC INSTITUTE, INC.

230 Park Avenue MUrray Hill 4-7800 New York 17, N. Y.

Telephone orders and COD's accepted

The Computing Laboratory in the University

Edited by Preston C. Hammer

Thirty-one experts from the 1955 conference on the computing laboratory in the university discuss the place of the university in the computing field, as well as particular applications of computers in science and industry, the problems of training personnel for computing, and ways to establish computing laboratories in universities.

256 pages $6.50

The University of Wisconsin Press

430 STERLING COURT • MADISON 6, WISCONSIN

Please mention the Journal of the AMERICAN STATISTICAL ASSOCIATION *in writing advertisers*

CHAPTER PRESIDENTS

ALBANY	Herbert L. Bryan, *Director of Research and Stat., New York State Dept. of Correction, 39 Columbia Street, Albany, New York*
AUSTIN	Tom H. Taylor, *Superintendent of Information & Stat., Texas Highway Department, Austin, Texas*
BOSTON	John G. Booton, *267 Beacon St., Boston, Massachusetts*
BUFFALO-NIAGARA	Robert H. Matthais, *20 Ward Park Road, Grand Island, New York*
CENTRAL INDIANA	Donald L. Cheak, *5432 East 19th St., Indianapolis, Indiana*
CENTRAL IOWA	Herbert T. David, *c/o Statistical Laboratory, Iowa State College, Ames, Iowa*
CENTRAL NEW JERSEY	Martin B. Wilk, *Bell Telephone Laboratories, Inc., Room 2D-535, Murray Hill, New Jersey*
CHICAGO	Rudolf T. Ericson, *4142 Howard Avenue, Western Springs, Illinois*
CINCINNATI	Harry Smith, Jr., *124 Hosea, Cincinnati 20, Ohio*
CLEVELAND	David Frazier, *Research Div., Standard Oil Company of Ohio, 2127 Cornell Rd., Cleveland 6, Ohio*
COLUMBUS	James C. Yocum, *Bureau of Business Research, The Ohio State University, Columbus, Ohio*
CONNECTICUT	Colin White, *81 Russell St., Hamden, Connecticut*
DAYTON	Richard E. Biedenbender, *208 Young St., Middletown, Ohio*
DENVER	Dennis Stump, Jr., *1205 S. Harrison Street, Denver 10, Colorado*
DETROIT	T. J. Obal, *Economics Analysis Dept., Ford Motor Co., Detroit, Michigan*
HAWAII	Charles F. Congdon, *College of Business Adm., University of Hawaii, Honolulu, Hawaii*
ILLINOIS	John A. Henry, *Dept. of Mechanical Engineering, University of Illinois, Urbana, Illinois*
ITHACA	C. R. Henderson, *Department of Husbandry, Cornell University, Ithaca, New York*
MILWAUKEE	J. Raymond Helbert, *Red Star Yeast and Products Co., 325 N. 27th St., Milwaukee 8, Wisconsin*
MONTREAL	Roger Lessard, *Ecole Polytechnique, 1430 St. Denis, Montreal, Que., Canada*
NEW ORLEANS	Roland Pertuit, *4871 Metropolitan Drive, New Orleans, Louisiana*
NEW YORK	Douglas Greenwald, *McGraw-Hill, 330 West 42nd Street, New York 36, New York*
NORTH CAROLINA	Gertrude M. Cox, *Institute o Statistics, Box 5457, State College Station, Raleigh, North Carolina*
NORTH TEXAS	Paul D. Minton, *Mathematics Dept., Southern Methodist University, Dallas, Texas*
OKLAHOMA CITY	William Wells, *Statistical Control Branch, Tinker Air Force Base, Midwest City, Oklahoma*
PHILADELPHIA	Haym Jaffe, *806 Brighton St., Philadelphia 11, Pennsylvania*
PITTSBURGH	C. C. Li, *School of Public Health, University of Pittsburgh, Pittsburgh 13, Pennsylvania*
PUERTO RICO	Juan Gonzales Ramos, *Puerto Rico Planning Board, Santurce, Puerto Rico*
ROCHESTER, N. Y.	Jack Karger, *210 East Hickory Street, East Rochester, New York*
SACRAMENTO	H. Arthur Domonoske, *5540 Spillman Avenue, Sacramento, California*
SAN FRANCISCO	Miss Phillis Beattie, *U. S. Bureau of Labor Statistics, 630 Sansome Street, Room 802, San Francisco 11, California*
SOUTHERN CALIFORNIA	John A. Scott, *1417 Oak St., Santa Monica, California*
STATE COLLEGE, PA.	John C. Griffiths, *Professor of Meteorology, Pennsylvania State University, University Park, Pennsylvania*
ST. LOUIS	Donald E. Paul, *c/o Laclede Gas Co., 1017 Olive St., St. Louis 1, Missouri*
TULSA	Lyndral E. Marcum, *Blue Cross Shield Hospital & Physician Service, P.O. Box 1738, Tulsa, Oklahoma*
TWIN CITIES (MINN.)	John Neter, *School of Business Adm., University of Minnesota, Minneapolis 14, Minnesota*
VIRGINIA	John W. Griswold, *810 Sunset Ave., W.H., Petersburg, Va.*
WASHINGTON, D.C.	Harold F. Dorn, *15 Burning Tree Court, Bethesda 14, Md.*